Cuk- H03685-S-PD24843

ANNUAL REVIEW OF MEDICINE:

Selected Topics in the Clinical Sciences

5

P, 24843

EDITORIAL COMMITTEE (1975)

C. H. COGGINS
Harvard Medical School

W. P. CREGER
Stanford University

E. W. HANCOCK
Stanford University

D. C. SABISTON JR.
Duke University

E. J. SACHAR
Albert Einstein College of Medicine

J. P. SANFORD
The University of Texas

M. SIPERSTEIN
University of California Service
Veteran's Administration Hospital,
San Francisco

R. T. SMITH
The University of Florida

L. W. WANNAMAKER
The University of Minnesota

Responsible for the organization of Volume 26
(Editorial Committee, 1973)

C. H. COGGINS
Harvard Medical School

W. P. CREGER
Stanford University

N. FREINKEL
Northwestern University

E. W. HANCOCK
Stanford University

S. S. KETY
Harvard Medical School

D. C. SABISTON JR.
Duke University

J. P. SANFORD
The University of Texas

R. T. SMITH
The University of Florida

L. W. WANNAMAKER
The University of Minnesota

Assistant Editors

J. CUTHBERTSON
L. RICHARDS

ANNUAL REVIEW OF MEDICINE:

Selected Topics in the Clinical Sciences

WILLIAM P. CREGER, *Editor*
Stanford University School of Medicine

C. H. COGGINS, *Associate Editor*
Harvard Medical School

E. W. HANCOCK, *Associate Editor*
Stanford University School of Medicine

VOLUME 26

1975

ANNUAL REVIEWS INC. 4139 EL CAMINO WAY PALO ALTO, CALIFORNIA 94306

UNIVERSITY COLLEGE OF MEDICINE
LIBRARY
• C. U. •

ANNUAL REVIEWS INC.
Palo Alto, California, USA

COPYRIGHT © 1975 BY ANNUAL REVIEWS INC.
ALL RIGHTS RESERVED

International Standard Book Number: 0-8243-0526-4
Library of Congress Catalog Card Number: A51-1659

Annual Reviews Inc. and the Editors of its publications assume no responsibility for the statements expressed by the contributors to this Review.

1634

REPRINTS

The conspicuous number aligned in the margin with the title of each article in this volume is a key for use in ordering reprints. Available reprints are priced at the uniform rate of $1 each postpaid. Effective January 1, 1975 the minimum acceptable reprint order is 10 reprints and/or $10.00, prepaid. A quantity discount is available.

PRINTED AND BOUND IN THE UNITED STATES OF AMERICA

CONTENTS

(Some articles are listed under more than one heading.)

ANESTHESIOLOGY:

Ketamine Anesthesia, *Charles F. Lanning and Merel H. Harmel* 137

Effects of Anesthesia on Intermediary Metabolism, *Edward A. Brunner, S. C. Cheng, and M. L. Berman* 391

CARDIOLOGY:

Sudden Death and the Pre-Hospital Phase of Acute Myocardial Infarction, *Wayne A. Wallace and Paul N. Yu* 1

Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis, *L. P. Le Quesne* 63

Left Ventricular Function in Acute and Chronic Coronary Artery Disease, *Charles E. Rackley and Richard O. Russell, Jr.* 105

Clinical Pharmacology of Digitalis Glycosides, *James E. Doherty and James J. Kane* 159

Mechanisms of Supraventricular Tachycardias, *Bruce N. Goldreyer* 219

Selective Distal Splenorenal Shunt for Bleeding Esophageal Varices, *Robert B. Smith III and W. Dean Warren* 229

Low Renin Hypertension, *J. Caulie Gunnells, Jr. and William L. McGuffin, Jr.* 259

The Steal Syndromes, *James F. Toole and Charles Patrick McGraw* 321

The Total Correction of Congenital Heart Disease, *Eoin Aberdeen* 451

Pathophysiology of Heart Block, *John F. Moran and Rolf M. Gunnar* 471

CLINICAL PHARMACOLOGY:

Tardive Dyskinesia, *George W. Paulson* 75

Clinical Pharmacology of Digitalis Glycosides, *James E. Doherty and James J. Kane* 159

Analgesic Abuse and Renal Disease, *Thomas Murray and Martin Goldberg* 537

DERMATOLOGY:

Skin Lipids and Acne, *John S. Strauss, Peter E. Pochi, and Donald T. Downing* 27

ENDOCRINOLOGY:

Metabolic Response to Infection, *William R. Beisel* 9

The Role of Hormone Receptors in the Action of Adrenal Steroids, *David Feldman* 83

Regulation of Albumin Metabolism, *Marcus A. Rothschild, Murray Oratz, and Sidney S. Schreiber* 91

Muscle Amino Acid Metabolism and Gluconeogenesis, *Neil B. Ruderman* 245

Low Renin Hypertension, *J. Caulie Gunnells, Jr. and William L. McGuffin, Jr.* 259

Endocrine Therapy of Breast Cancer, *William Leo McGuire* 353

Interconversions of Apolipoprotein Fragments, *Conrad B. Blum and Robert I. Levy* 365

Effects of Anesthesia on Intermediary Metabolism, *Edward A. Brunner, S. C. Cheng, and M. L. Berman* 391

Gonadotropin-Releasing Hormone, *Samuel S. C. Yen* 403

Relation Between Growth Hormone and Somatomedin, *Judson J. Van Wyk and Louis E. Underwood* 427

Active Form of the Thyroid Hormone and Somatomedin, *Sidney H. Ingbar and Lewis E. Braverman* 443

Hyperglycemia, Polyol Metabolism, and the Complications of Diabetes Mellitus, *Kenneth H. Gabbay* 521

GASTROENTEROLOGY:

Regulation of Albumin Metabolism, *Marcus A. Rothschild, Murray Oratz, and Sidney S. Schreiber* 91

Pathogenesis of Hepatic Coma, *Leslie Zieve and Demetre M. Nicoloff* 143

Selective Distal Splenorenal Shunt for Bleeding Esophageal Varices, *Robert B. Smith III and W. Dean Warren* 229

Disorders of the Lower Esophageal Sphincter, *Robert S. Fisher and Sidney Cohen* 373

HEMATOLOGY:

Von Willebrand's Disease, *Lars Holmberg and Inga Marie Nilsson* 33

Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis, *L. P. LeQuesne* 63

Cellular Kinetics, Staging, and Immunoglobulin Synthesis in Multiple Myeloma, *Brian G. M. Durie and Sydney E. Salmon* 283

Granulocyte Transfusion Therapy, *Donald J. Higby and Edward S. Henderson* 289

The Antiglobulin Test in Autoimmune Hemolytic Anemia, *Wendell F. Rosse* 331

Thalassemia, *Bernard G. Forget and David G. Nathan* 345

IMMUNOLOGY:

Alterations in T- and B-Cells in Human Disease States, *Ralph C. Williams, Jr. and Ronald P. Messner* 181

The Antiglobulin Test in Autoimmune Hemolytic Anemia, *Wendell F. Rosse* 331

Immunotherapy (Desensitization) in Allergic Disease, *Philip S. Norman* 337

INFECTIOUS DISEASE:

Metabolic Response to Infection, *William R. Beisel* 9

Granulocyte Transfusion Therapy, *Donald J. Higby and Edward S. Henderson* 289

Urinary Tract Symptoms and Infections, *Jay P. Sanford* 485

MISCELLANEOUS:

Dental Caries, *R. J. Gibbons and J. van Houte* 121

Unusual Etiologies of Fever and Their Evaluation, *Sheldon M. Wolff, Anthony S. Fauci, and David C. Dale* 277

NEOPLASIA:

Scanning Techniques to Assess Thyroid Nodules, *James L. Quinn, III and Robert E. Henkin* 193

Treatment of Testicular Cancer, *Donald I. Twito and B. J. Kennedy* 235

Cellular Kinetics, Staging, and Immunoglobulin Synthesis in Multiple Myeloma, *Brian G. M. Durie and Sydney E. Salmon* 283

Endocrine Therapy of Breast Cancer, *William Leo McGuire* 353

The Role of Radiation Therapy in the Definitive Treatment of Adenocarcinoma of the Prostate, *Gordon R. Ray and Malcolm A. Bagshaw* 567

NEPHROLOGY:

Calcium Oxalate Renal Stones, *Edwin L. Prien, Jr.* 173

Low Renin Hypertension, *J. Caulie Gunnells, Jr. and William L. McGuffin, Jr.* 259

Acute Renal Failure (Vasomotor Nephropathy): Micropunture Studies of the Pathogenetic Mechanism, *Donald E. Oken* 307

Urinary Tract Symptoms and Infections, *Jay P. Sanford* 485

Analgesic Abuse and Renal Disease, *Thomas Murray and Martin Goldberg* 537

NEUROLOGY:

The Minimal Brain Dysfunction Syndrome, *Paul H. Wender* 45

Tardive Dyskinesia, *George W. Paulson* 75

Pathogenesis of Hepatic Coma, *Leslie Zieve and Demetre M. Nicoloff* 143

The Steal Syndromes, *James F. Toole and Charles Patrick McGraw* 321

Huntington's Disease, *Ira Shoulson and Thomas N. Chase* 419

A Definition of Cerebral Death, *Benjamin Boshes* 465

Neuroregulators and Sleep Mechanisms, *R. Bruce Holman, Glen R. Elliott, and Jack D. Barchas* 499

Subacute Sclerosing Panencephalitis, *Richard S. Johannes and John L. Sever* 589

PSYCHIATRY:

X-Linkage in Manic-Depressive Illness, *Remi J. Cadoret and George Winokur* 21

Biochemical Aspects of Schizophrenia, *Steven Matthysse and Joseph Lipinski* 551

RADIOLOGY:

Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis, *L. P. Le Quesne* 63

Scanning Techniques to Assess Thyroid Nodules, *James L. Quinn, III and Robert E. Henkin* 193

The Role of Radiation Therapy in the Definitive Treatment of Adenocarcinoma of the Prostate, *Gordon R. Ray and Malcolm A. Bagshaw* 567

REPRODUCTIVE SYSTEM:

Gonadotropin-Releasing Hormone, *Samuel S. C. Yen* 403

RESPIRATORY SYSTEM:

Oxygen Toxicity, *Walter G. Wolfe and William C. DeVries* 203

CUMULATIVE INDEX OF CONTRIBUTING AUTHORS, VOLUMES 22 TO 26 603

CUMULATIVE INDEX OF CHAPTER TITLES, VOLUMES 22 TO 26 605

ANNUAL REVIEWS INC. is a nonprofit corporation established to promote the advancement of the sciences. Beginning in 1932 with the *Annual Review of Biochemistry,* the Company has pursued as its principal function the publication of high quality, reasonably priced Annual Review volumes. The volumes are organized by Editors and Editorial Committees who invite qualified authors to contribute critical articles reviewing significant developments within each major discipline.

Annual Reviews Inc. is administered by a Board of Directors whose members serve without compensation.

BOARD OF DIRECTORS
1975

Dr. J. Murray Luck
Founder Emeritus, Annual Reviews Inc.
Department of Chemistry
Stanford University

Dr. Esmond E. Snell
President, Annual Reviews Inc.
Department of Biochemistry
University of California, Berkeley

Dr. Joshua Lederberg
Vice President, Annual Reviews Inc.
Department of Genetics
Stanford University Medical School

Dr. William O. Baker
President
Bell Telephone Laboratories

Dr. Ivan L. Bennett, Jr.
Provost and Dean
New York University Medical Center

Dr. James E. Howell
Graduate School of Business
Stanford University

Dr. Wolfgang K. H. Panofsky
Director
Stanford Linear Accelerator Center

Dr. John Pappenheimer
Department of Physiology
Harvard Medical School

Dr. Colin S. Pittendrigh
Department of Biological Sciences
Stanford University

Dr. Richard N. Thomas
Institut d'Astrophysique
Paris, France

Dr. Alvin M. Weinberg
Director, Energy Research and
Development Office
Federal Energy Bureau

Dr. Harriet Zuckerman
Department of Sociology
Columbia University

Annual Reviews are published in the following sciences: Anthropology, Astronomy and Astrophysics, Biochemistry, Biophysics and Bioengineering, Earth and Planetary Sciences, Ecology and Systematics, Entomology, Fluid Mechanics, Genetics, Materials Science, Medicine, Microbiology, Nuclear Science, Pharmacology, Physical Chemistry, Physiology, Phytopathology, Plant Physiology, Psychology, and Sociology. The *Annual Review of Energy* will begin publication in 1976. In addition, two special volumes have been published by Annual Reviews Inc.: *History of Entomology* (1973) and *The Excitement and Fascination of Science* (1965).

Copyright 1975. All rights reserved

SUDDEN DEATH AND THE PRE-HOSPITAL PHASE OF ACUTE MYOCARDIAL INFARCTION

◆7125

Wayne A. Wallace, M.D. and Paul N. Yu, M.D.[1]
Department of Medicine, University of Rochester School of Medicine and Dentistry,
Rochester, New York 14642

INTRODUCTION

Coronary artery disease is the leading cause of death in the United States, with nearly 670,000 fatalities reported in 1969 (1).

In 1962 the coronary care unit (CCU) was introduced as a special area in the hospital where acute coronary patients could be kept under constant electronic and nursing surveillance. Since that time there has been much emphasis and progress in the in-hospital care of patients with acute myocardial infarction, and the in-hospital mortality rate has decreased from an average of about 35% to 20%. However, numerous studies of coronary artery disease have reported that at least half of the deaths from acute myocardial infarction or coronary attack[2] occur outside hospitals (2–9). In the past few years new attention has been directed toward the problem of sudden death and the pre-hospital phase of acute myocardial infarction, as it appears that this is the area where further reductions in the overall mortality rate from acute myocardial infarction or coronary attack are to be made in the absence of a major breakthrough in the primary prevention of coronary atherosclerosis.

PRE-HOSPITAL DEATHS

Magnitude of the Problem

Epidemiologic studies have shown that approximately two thirds of the deaths from acute myocardial infarction or coronary attack occur in the pre-hospital phase.

---

[1] Supported in part by Contract NO1HV81331 (MIRU) from the Clinical Cardiac Diseases Branch, National Heart and Lung Institute.

[2] The term acute coronary attack is used here because in a certain number of patients who die suddenly no pathological evidence of acute myocardial infarction is detected in routine postmortem examination.

Table 1 Mortality following onset of symptoms of acute coronary events

Author	Percentage of deaths in 1 hr	Percentage of deaths in 2 hr	Percentage of deaths in 4 hr	Percentage of deaths before entering a hospital or seen by a physician
Bainton (2)	63			
Kuller 1966 (3)				67
Kuller 1972 (4)		63		
McNeilly (5)				60
Fulton (6)			55	
Feinleib (7)				60
Simon (8)	64		80	
Kinlen (9)	50			65

From 50 to 64% of the deaths occur within 1 hr of the onset of symptoms and 55 to 80% occur within 4 hr. The data from eight studies are summarized in Table 1. Fulton, Julian & Oliver (6) plotted the cumulative percentage mortality in the first 24 hr after the onset of symptoms from studies done in Scotland and Ireland. The data show an exponential curve with a very rapid rise in the first 2 hr and a slow rise in cumulative mortality after 6–8 hr. Simon & Alonzo (8) obtained a similar exponential cumulative mortality curve from a study of a community in the United States.

Mechanism of Death

The mechanism of sudden and unexpected death from coronary artery disease is most often ventricular fibrillation. Lawrie and co-workers (10) found in their experience that 13 of 15 episodes of primary ventricular fibrillation occurred within 5 hr of the onset of symptoms. Similarly, with the advent of mobile coronary care units, Pantridge reported that when patients were seen promptly, ventricular fibrillation was usually the cause of collapse (11). Of 87 patients seen within 4 min of collapse, 73 had ventricular fibrillation. The remaining 14 had ventricular asystole. However, in patients seen later, 80% had ventricular asystole. Cobb and co-workers (12) also observed that 54 of 76 patients (71%) with sudden, unexpected circulatory arrest had ventricular fibrillation. These observations were confirmed by the recent study of Liberthson and associates who reported that ventricular fibrillation was recorded in the majority of patients with sudden cardiac death when electrocardiographic monitoring was available (13).

Reversibility

Is this ventricular fibrillation which causes so many sudden out-of-hospital deaths a reversible disorder which can be treated with cardiopulmonary resuscitation and defibrillation? Early experiences in treating ventricular fibrillation in CCUs indicated that indeed this lethal cardiac arrhythmia can be successfully treated with long-term survival. Pantridge & Geddes (14) found that their CCU patients treated for ventricular fibrillation had a 62% survival rate. When primary ventricular fibrillation is analyzed separately from ventricular fibrillation secondary to congestive heart failure, cardiogenic shock, cardiac drugs, intracardiac pacemakers, etc, a higher percentage of survival is observed. Lawrie and co-workers (10) reported that in their CCU 87% of patients with primary ventricular fibrillation survived. Mounsey (15) likewise found that in 17 of 18 patients primary ventricular fibrillation could successfully be converted to sinus rhythm. Long-term survival is also excellent. Stannard & Sloman (16) reported that 18 of 20 patients resuscitated from primary ventricular fibrillation survived, with the longest survival 4½ yr. Geddes and co-workers (17) also found that survivors of ventricular fibrillation had a long-term prognosis similar to that of patients with myocardial infarction not complicated by ventricular fibrillation.

PRE-HOSPITAL EMERGENCY CARE OF ACUTE MYOCARDIAL INFARCTION

The pre-hospital emergency care of acute coronary events is an approachable problem since the mechanism of sudden cardiac death is now understood to be predominantly ventricular fibrillation and since this arrhythmia can most often be completely reversed if appropriate emergency medical facilities are available.

Delays in Application of Early Medical Care

The first problem to be considered is the delay in application of early medical care. The major causes of this delay include the patients, physicians, transportation system, and receiving areas. Several workers have analyzed the components of delay that usually occur between the onset of symptoms and the arrival at a medical facility by traditional health-care systems. In several studies this delay has generally been 4 hr or more (5, 6, 8a, 10), with patient's decision time representing the largest portion. Simon and co-workers (8a) found that 24% of the patients stated that they initially considered calling a physician earlier in their symptomatic phase but elected not to because of the hour of the day, the presumed unavailability of the physician, or because their interpretation of the severity of the symptoms did not dictate medical consultation. Also, Olin & Hackett (18) reported that much of the patient delay is apparently due to denial of the significance of symptoms; they found that 9 of 32 patients who correctly diagnosed the cause of their chest pain failed to take appropriate steps to obtain treatment. They also noted that a history of myocardial infarction or angina pectoris failed to cause patients to seek help sooner.

The delay in early medical care due to physicians' indecision or unavailability has appreciably decreased. Most physicians now instruct their patients to go to the emergency department of a hospital or call a mobile coronary care unit whenever they have severe retrosternal chest pain lasting more than 15 or 20 min.

Various analyses of the delay caused by transporting patients with acute myocardial infarction reveal that this contributes a relatively small fraction to the total delay. Fulton (6) found that the ambulance time accounted for 10% of the delay from onset of symptoms to arrival at a CCU. Simon and co-workers (8a) found that transportation time accounted for 11% of the time to definitive care. Feinleib & Davidson (7) report a travel time of 12 min out of 4.3 hr, or 5% of the total time.

The fourth component of delay in early medical care is the receiving area. For example, it is possible or quite likely that in the emergency department of a hospital a patient with symptoms suggestive of acute myocardial infarction may not get as immediate attention as one who has fracture of the leg or hemoptysis. The establishment of mobile care units has largely circumvented this delay.

Approaches to Early Application of Emergency Care

In 1971 the Inter-Society Commission for Heart Disease Resources (19) outlined and recommended a system to shorten the delay in providing medical care to patients with suspected or proven acute myocardial infarction. Their recommendations included professional and public educational programs, mechanisms to bring patients to early medical care, and establishment of life-support stations to receive these patients. Two major types of life-support stations were proposed: fixed and mobile. The concept of life-support stations has been supported by several workers, and all have recognized the need for training paramedical personnel in the emergency care of coronary events and the necessity for having a system capable of monitoring and stabilizing cardiac rhythm at the earliest possible moment. Lown & Ruberman (20) and Yu (21, 22) have suggested pre-coronary areas in emergency rooms, which would provide ambulatory triage and encourage self-referral. Such fixed life-support stations in emergency rooms of hospitals would be appropriate for both large and small communities and hospitals throughout the country, whereas mobile life-support units may be feasible and suitable in a limited number of large cities.

The concept of mobile coronary care units was first introduced by Pantridge and associates of Belfast, Ireland in 1966 (13). Their early reports (23, 24) indicated that they could reach about 78% of the patients within 15 min. They saw 48% of their patients within 2 hr of onset of symptoms. In 1970 (11) they reported that the median delay from onset of symptoms to arrival of the mobile coronary care unit was about 1 hr 40 min. In patients under 70 in their series, the hospital mortality rate was approximately 10.5%. In patients under 70 who were seen within 1 hr of the onset of symptoms, the hospital mortality rate was 8.6% and the incidence of cardiogenic shock was 3.4% In the United States, Grace & Chadbourn (25) found that in 50 patients attended by the mobile coronary care unit there were only three deaths or a 6% hospital mortality. In 1974, Yu (26) reported available data from mobile coronary care units or mobile intensive care units in 16 American cities, which showed that about 20% of the patients with ventricular fibrillation who were successfully resuscitated left the hospital alive.

More recently data have been collected from fixed life-support stations (27, 28). Kassanoff and co-workers (27) have reported on a fixed life-support station, which they established in a large stadium. They found an apparently low incidence of sudden cardiac deaths in the stadium population and concluded that it was an impractical endeavor in terms of both cost and personnel. On the other hand, since 1972 the authors (29) have been caring for patients in a fixed life-support station established in St. Mary's Hospital in Rochester, New York. All adult patients with chest pain entering the emergency department are immediately placed in a special two-bed unit staffed with a trained nurse and with monitoring and resuscitation equipment. Between January 3, 1972 and September 30, 1973, 1064 patients were screened in this station. Of these, 154 had documented acute myocardial infarction which had occurred within 24 hr of presenting themselves to the station. The median arrival time was 66 min in 43% of the patients seen within 2 hr of their symptomatic onset, which compares favorably with the experience of mobile coronary care units. The overall hospital mortality rate was 15.6%. However, in patients under 70 years of age, the hospital mortality rate was 7%. The hospital mortality rate of patients under 70 years who were seen within 1 hr after symptomatic onset was only 4%. The incidence of cardiogenic shock and of left ventricular failure was also lower in patients who arrived within 2 hr than in those who arrived later.

Emergency Treatment

In addition to the capability for defibrillation in the life-support station, it is also important that facilities be available to initiate antiarrhythmic therapy. Generally, intravenous administration of lidocaine in CCUs has been effective in eradicating ventricular premature contractions and preventing ventricular fibrillation. There is some evidence presented by Adgey & Pantridge (30), however, that lidocaine may not be as effective in the first 2 hr after acute myocardial infarction. In 66 of their patients seen within 2 hr of the onset of symptoms, they found that ventricular ectopics were abolished in 33%, reduced in 38%, and not affected in 29%. Pantridge (31) reports that procainamide is effective in the very early phase of disease, particularly in patients with heart rates exceeding 90 per min. Therefore, in spite of the general effectiveness of intravenous lidocaine, further study is required to see if procainamide or some other antiarrhythmic agent may be more efficacious to control ventricular ectopics in this very early phase of acute coronary events.

Intravenous administration of atropine is a widely accepted therapy for sinus bradycardia associated with hypotension in acute myocardial infarction. Several workers have reported beneficial effects of intravenous atropine on cardiac rhythm and a decrease in mortality in this setting (24, 30–32). Occasionally, however, ventricular ectopics or even ventricular fibrillation have appeared for the first time after the administration of atropine. Recently, Morris, Mercer & Yeates (33) reported observations on 735 patients admitted to a CCU. They found a lower hospital mortality rate in patients with sinus bradycardia (6%) than in patients with normal heart rate (15%) or in patients with sinus tachycardia (26%). However, as they point out, sinus bradycardia at the onset of myocardial infarction may behave differently from sinus bradycardia seen later in the course of the disease. Epstein and co-workers found that atropine did not decrease malignant ventricular premature contractions or ventricular tachycardia in dogs (34). They also reported that a heart rate above 80 per min after atropine administration resulted in an increase in serious ventricular arrhythmias and caused an increase in the vulnerability to ventricular fibrillation. Whether these findings are significant for patients with acute myocardial infarction is not yet known and further investigation is required.

CONCLUDING REMARKS

Following the success of hospital CCUs it has been increasingly recognized that if further reduction in the mortality of acute myocardial infarction or coronary attack is to be realized, it is imperative to deliver medical care to patients very soon after their acute coronary event. This is rationalized because the majority of deaths occur in the first 2–4 hr after onset of symptoms, or before a patient reaches medical facilities in the traditional health care delivery systems. Since the majority of these early and sudden deaths are due to ventricular fibrillation and since this is a readily reversible arrhythmia with a good long-term survival, the outlook is rather promising. The major cause of delay in the delivery of early medical care has been identified as patient delay, so public and physician education is of paramount importance. Once the significance of symptoms is appreciated, the patient must have very rapid access to medical facilities capable of monitoring the heart rhythm, reversing ventricular fibrillation, and administering appropriate antiarrhythmic therapy. For larger cities, mobile coronary care units are one approach to this problem. Reported experiences from Europe and the United States demonstrate that this method can be very effective in reaching patients early, in resuscitating effectively, and in decreasing the overall mortality. Our experience indicates that a fixed life-support station established in a community hospital can also result in the very early application of cardiac care and likewise reduce mortality. Fixed life-support stations may be more feasible and practical for the majority of urban and rural areas.

Literature Cited

1. *Deaths Due to Cardiovascular Diseases by Major Types of Disorder in United States.* 1969. Washington, D.C: Nat. Center Health Statist., USPHS, Dept. HEW
2. Bainton, C. R., Peterson, D. R. 1963. *N. Engl. J. Med.* 268:569–75
3. Kuller, L., Lilienfeld, A., Fisher, R. 1966. *Circulation* 34:1056–68
4. Kuller, L., Cooper, M., Perper, J. 1972. *Arch. Intern. Med.* 129:714–19
5. McNeilly, R. H., Pemberton, J. 1968. *Brit. Med. J.* 3:139–42
6. Fulton, M., Julian, D. G., Oliver, M. F. 1969. *Circulation* 40 (Suppl. IV): 182–91
7. Feinleib, J., Davidson, J. J. 1972. *J. Am. Med. Assoc.* 222:1129–34
8. Simon, A. B, Alonzo, A. A. 1973. *Arch. Intern. Med.* 132:163–70

8a. Simon, A. B., Feinleib, M., Thompson, H. K. *Am. J. Cardiol.* 30: 476–82

9. Kinlen, L. J. 1973. *Brit. Heart J.* 35:616–22
10. Lawrie, D. M. et al 1967. *Lancet* 2:109–14
11. Pantridge, J. F. 1970. *Chest* 58:229–34
12. Cobb, L. A., Conn, R. D., Samson, W. E., Philbin, J. E. 1970. *Circulation* 42 (Suppl. III):144 (Abstr.)
13. Liberthson, R. R. et al 1974. *Circulation* 49:790–98
14. Pantridge, J. F., Geddes, J. S. 1966. *Lancet* 1:807–8
15. Mounsey, P. 1967. *Am. J. Cardiol.* 20:475–83
16. Stannard, M., Sloman, G. 1969. *Am. Heart J.* 77:573
17. Geddes, J. S., Adgey, A. A. J., Pantridge, J. F. 1967. *Lancet* 2:273–75
18. Olin, J. S., Hackett, T. P. 1964. *J. Am. Med. Assoc.* 190:977–81
19. Study Group on Coronary Artery Disease. 1971. Report of the inter-society commission for heart disease resources. *Circulation* 43*A*:171–83
20. Lown, B., Ruberman, W. 1970. *Mod. Conc. Cardiovasc. Dis.* 39:97–102
21. Yu, P. N. 1971. *Circulation* 44:979–81
22. Yu, P. N. 1972. *Circulation* 45:189–204
23. Pantridge, J. F., Geddes, J. S. 1967. *Lancet* 2:271–73
24. Adgey, A. A. J., Geddes, J. S., Mulholland, H. C., Keegan, D. A. J., Pantridge, J. F. 1968. *Lancet* 2:1097–1101

25. Grace, W. J., Chadbourn, J. A. 1970. *Circulation* 42 (Suppl. III):160 (Abstr.)
26. Yu, P. N. 1974. *Arch. Intern. Med.* 134:234–39
27. Kassanoff, I. et al 1972. *J. Am. Med. Assoc.* 221:397–99
28. Carveth, S. W. 1968. *Dis. Chest* 53:8–11
29. Napodano, R. J., Wallace, W. A., Vargas, A. A., Yu, P. N. 1972. *Circulation* 46 (Suppl. II):197 (Abstr.)
30. Adgey, A. A. J., Pantridge, J. F. 1972. *Geriatrics* 27:102–10
31. Pantridge, J. F. 1974. *Pre-Hospital Management of Acute Ischemic Heart Disease.* Panel discussion at 23rd Ann. Sci. Sess., Am. Coll. Cardiol., New York, NY
32. Lewis, R. P., Warren, J. V. 1974. *Am. J. Cardiol.* 33:151 (Abstr.)
33. Morris, R. M., Mercer, C. J., Yeates, S. E. 1972. *Brit. Heart J.* 34:901–4
34. Epstein, S. E., Goldstein, R. E., Redwood, D. R., Kent, K. M., Smith, E. R. 1973. *Ann. Intern. Med.* 78:918–36

Copyright 1975. All rights reserved

METABOLIC RESPONSE TO INFECTION

❖7126

William R. Beisel, M.D.
U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, Maryland 21701

INTRODUCTION

An invasion of pathogenic microorganisms causes the body to initiate a variety of defensive reactions. Responses such as the mobilization of phagocytes, the development of localized inflammation, and the activation of immunogenic mechanisms have relatively distinct, identifiable roles in host defense. Underlying and apparently in support of such purposeful responses are a number of poorly understood, generalized host metabolic responses without, at present, a known or well-defined defensive function. Nevertheless, the very extent and complexity of such metabolic responses in the infected host demand that they be recognized and defined in terms of their potential contributions to resistance and survival. These changes should not be dismissed conceptually as events that occur only in a random or haphazard fashion. Rather, certain patterns of metabolic response are highly consistent in their occurrence and can be documented in many infectious illnesses (1).

Some of the catabolic responses during a febrile infection are quite stereotyped in pattern and evolve in a sequence that is repetitive, predictable, and analogous to catabolic responses characteristic of trauma (1, 2). Other consistent metabolic changes are initiated by endogenous signal mechanisms triggered during phagocytic activity. Still other types of host metabolic response seem to develop during certain infections only (3).

GENERAL CONCEPTS

Despite many complexities, currently available data allow the formulation of broad concepts for interpreting biochemical and endocrine changes during infection:

(*a*) Metabolic responses begin promptly following the initiation of an infectious process and continue to evolve as a series of interrelated events. These responses are geared in their timing and magnitude to the clinical stages and severity of illness. As will be reviewed, many host biochemical changes begin within hours of the invasion of an infectious microorganism, well before the onset of any clinical sign or symptom of illness (1, 4, 5). Some subsequent metabolic responses coincide with the onset of fever, and still others accompany the onset of convalescence. Persistence of individual metabolic changes is also variable following recovery.

(*b*) Metabolic changes in the host may be initiated by the direct effect on body cells of a microorganism or its toxic products. Other responses occur as a secondary consequence of the inflammatory, febrile, neural, cardiovascular, immunological, or nutritional manifestations of an infectious process.

(*c*) Metabolic responses may exhibit biphasic patterns of sequential change. Thus, accurate analytical data if gathered at only a single point in time during an infection may give rise to a distorted interpretation. For example, hyperglycemia and glucose intolerance during the early stages of an infectious process (6) may give way to hypoglycemia and a depletion of tissue carbohydrate in an uncontrolled infection (7–9).

(*d*) Metabolic responses of the host are clearly influenced by the severity and duration of illness, by the effects of therapy, by preexisting factors in the genetic, immunological, and nutritional status of the host, and even by the time of day that an infectious process is initiated.

(*e*) Although a generalized infectious process may initiate a characteristic group of metabolic sequelae, these will be modified by the concomitant development of unique pathogenic events such as liver cell dysfunction during viral hepatitis, muscle paralysis during poliomyelitis, progressive hypotension during gram-negative sepsis, or electrolyte depletion during cholera-like diarrhea.

(*f*) Because multiple variables are involved in the broad spectrum of metabolic responses to infection, special care must be taken when interpreting specific observations. Changes in concentration of a substance in body pools such as blood, plasma, or tissue must be evaluated with due consideration given to possible alterations of pool size as well as to kinetic rates of entry and egress of the substance at the time it is measured within the pool. Thus, a mechanistic explanation for an observed change, such as agonal hypoglycemia, demands that data be obtained relative to the rates of glucose production and utilization, the activity of carbohydrate-regulating hormones and cellular enzymes, as well as the availability of suitable substrates.

GENERALIZED CATABOLIC RESPONSES

Clinically the most prominent metabolic response during a febrile infection is the wasting of body tissues (1, 2, 10). Although anabolic responses also occur during a generalized infection, they are generally overshadowed in magnitude by the greater clinical prominence of catabolic events. Nitrogen balance becomes negative despite the fact that both anabolic and catabolic processes are occurring simultaneously during the febrile periods of infection. The catabolism of muscle protein serves to redistribute amino acids required for the synthesis of protein in other locations. Thus, the body seems able to sacrifice large quantities of muscle protein in order to provide precursor materials needed for meeting both the anabolic requirements for key host defensive responses and the heightened requirements for metabolizable energy (11).

Fever appears to be the major stimulus for initiating catabolic losses during an acute infectious illness (12). Widespread catabolism of body tissues does not begin until after fever has become fully evident. A febrile response requires the excessive expenditure of body energy even though fever is generally accompanied by a diminished intake of food and other nutrients. In combination with continued or increased excretory losses through urine, stool, and sweat, these changes lead inevitably to negative body balances of many substances.

Deficits in Metabolic Balances

Losses of body nitrogen serve as the most widely studied example of the catabolic response to infection, but the other major elements found in intracellular fluids (K, P, Mg, S, and Zn) all appear to be lost in quantities roughly proportional to those of nitrogen (2). These losses plus increased expenditures of metabolic energy cause body weight and muscle mass to decline. Since the body does not store large amounts of free amino acids, a negative balance for nitrogen implies a net loss of body protein during the febrile phase of an infectious process. Negative balances result from sustained or increased excretory loss of nitrogen in the face of a reduction in dietary nitrogen intake. If a normal person lowers or temporarily discontinues his intake of dietary protein, the body responds quickly by reducing urinary nitrogen losses (13). This compensatory response does not generally occur during febrile infections. Rather, urinary losses of nitrogenous compounds such as urea, ammonia, creatinine, and amino acids tend to be increased or maintained at close to normal rates. Nitrogen losses in the stool do not generally increase unless diarrhea is present.

Despite the fact that most catabolic responses are relatively stereotyped, unusual losses of specific nitrogenous compounds in urine serve in some instances to identify an altered metabolic function in the infected host. For example, the increased excretion of uric acid during infectious mononucleosis is thought to signify an increased turnover of purines (14).

If a fever is induced artificially in normal volunteers, negative metabolic balances develop rapidly and are of a magnitude comparable to those measured during febrile infections (12). In contrast, neither the restriction of dietary intake nor the administration of exogenous glucocorticoid hormones to noninfected volunteers (when done so as to mimic the observed patterns of host response during infection) cause changes in metabolic balances characteristic in timing or magnitude of those during acute infection (2).

The severity of a brief fever appears to correlate with the magnitude of catabolic losses. However, should an infectious process become subacute or chronic, daily balances become progressively less negative and eventually the patient enters a chronic state of cachexia such as that seen typically in patients with advanced tuberculosis. During convalescence from an infectious process, negative balances are reversed. Depleted body stores can be regained, but the reconstitution of body deficits following infection is rather slow, and like the prolonged metabolic convalescence from trauma or surgery, may take many weeks (1, 2). If the nutritional status of a child is marginal the nitrogen losses associated with an infection such as measles or weanling diarrhea may precipitate acute kwashiorkor (10).

Degradation of Protein

Several points of evidence suggest that the degradation of skeletal muscle protein is accelerated during periods of infection. In vivo catabolism of contractile protein in man can be estimated by measurements of 3-methylhistidine in urine. This analogue of histidine is created in situ and in defined molar quantities when the contractile muscle proteins actin and myosin are synthesized (13). Since this unique amino acid is subjected to little if any degradation to CO_2 or reutilization in man, its excretion in the urine can serve to quantitate the daily catabolism of contractile proteins. Thus, serial measurements of urinary 3-methylhistidine excretion were used by Wannemacher et al (15) to compare turnover rates of contractile proteins during health and during a period of acute infection. These studies provided data in close agreement with estimates derived from increased urinary creatine and creatinine excretion, negative nitrogen balances, and body weight losses, supporting the view that skeletal muscle catabolism is a major metabolic response during a febrile infection. Mechanisms that initiate the changes in muscle catabolism have not been identified, although infection-induced alterations in the plasma glucagon: insulin ratio may be involved.

GENERAL ANABOLIC RESPONSES

Unlike the stereotyped pattern of catabolic wasting during infection, the anabolic responses appear to vary in form, timing, or magnitude in relation to the causative microorganism. Anabolic responses may be initiated very early in an infectious process; they persist despite the subsequent appearance of catabolic phenomena (1, 2).

The generation of phagocytic cells and the synthesis of specific immunoglobulins, as well as the production of components of the complement system, interferon, acute phase serum glycoglobulins, and a variety of intracellular enzymes, metabolic products and mediators, all require the availability of adequate quantities of cellular energy as well as intact protein synthesizing mechanisms. Indeed, every aspect of host defense against infection is ultimately dependent upon the ability of host cells to synthesize protein.

Infection-related changes in the metabolism of individual cells with a specific role in host defense provide some insight into the molecular mechanisms needed to support these activities. During phagocytosis there is a burst of metabolic energy within the cell, the formation of phagocytic vacuoles, and activation and release of many lysosomal enzymes (16, 17). Other examples include the biochemical mechanisms used to reorient metabolic systems of host cells for viral replication (18), and the metabolic changes within macrophage (19) and lymphoid cell populations (20) when they are involved in the activation of an immune response.

In addition to the specific anabolic responses of disparate cells, protein synthesis is accelerated in the liver and other visceral tissues during infection. There is an increased activity of intracellular hepatic enzymes such as tyrosine transaminase and tryptophan oxygenase (21), and of the hepatic production of acute phase reactant serum glycoproteins (22) including haptoglobin, α_1-antitrypsin, α_1-acid glycoprotein, fibrinogen, C-reactive protein, ceruloplasmin, and seromucoid. Although the functions of these proteins have yet to be elucidated in terms of their contributions to host defenses, protein synthesis of this type occurs even during periods of starvation or extreme protein-calorie malnutrition (11).

ENDOGENOUS MEDIATORS

An important new concept in understanding generalized metabolic sequelae during infection or other inflammatory reactions arises from the observation that the body utilizes a hormone-like signal mechanism to initiate nonspecific metabolic responses in the liver and other anatomic locations. This mechanism involves the release of endogenous protein messengers by phagocytizing cells which have been activated by the uptake of infectious microorganisms or the stimulation of substances such as endotoxin, interferon inducers, phlogistic agents, or the products of antigen-sensitized lymphocytes (see Figure 1). Unlike the regional action of cellular products that initiate localized inflammatory responses, mediators such as endogenous pyrogens (EP) or leukocytic endogenous mediator (LEM) travel via the blood to act on distant sites following their release from activated phagocytic cells.

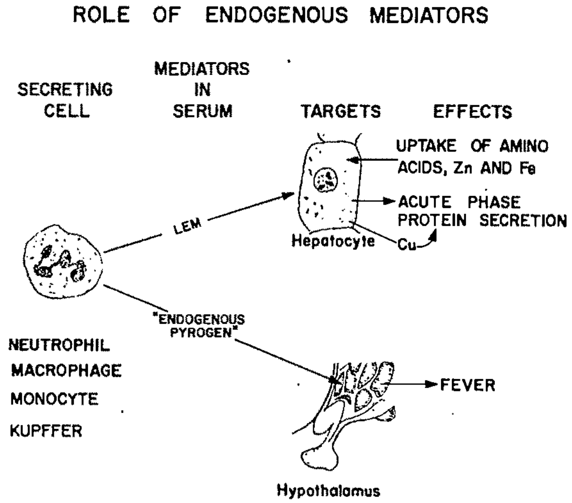


Figure 1. In addition to their phagocytic capabilities, certain body cells act as fixed or mobile sentinels with the capability of releasing hormone-like endogenous mediators as signals to initiate generalized host responses in distant tissues.

The ability of phagocytizing cells to initiate a febrile response by releasing EP is well recognized (24). In like manner, a similar or perhaps identical substance, LEM, has been shown to initiate many types of host response (25, 26). LEM acts upon the liver rather than on the hypothalamus, stimulating both a hepatic uptake of amino acids, zinc, and iron from the plasma, as well as an increased hepatic synthesis of acute phase reactant glycoproteins. LEM also appears to serve as a signal that initiates an increased release of granulocytes from the marrow (26).

Using a sensitive bioassay technique, LEM has been shown to circulate in the plasma of febrile patients with many different kinds of acute infections (4, 25). LEM is readily found in the plasma of experimentally infected laboratory animals and can be produced in vitro by phagocytizing neutrophils, monocytes, or macrophages. Although LEM has not as yet been isolated in pure form, it appears to be a heat-labile, trypsin and pronase-sensitive protein with a molecular weight of 10,000 to 30,000. It is stable on freezing or lyophilization and retains its biological activity after exposure to lipase, RNAse, or a wide range of pH (25, 26).

The LEM-induced stimulation of amino acid influx into the liver is known to involve both the A- and L-type membrane transport sites and is accomplished without activation of adenylate cyclase, cyclic AMP formation, or the participation of other hormones (26). The ability of LEM to initiate a massive synthesis of acute phase globulins on the rough endoplasmic reticulum begins with an initial activation of templates of nuclear DNA chromatin (29) followed in turn by the synthesis of RNA and protein. This latter action seems to require the permissive presence of physiologic concentrations of the adrenal glucocorticoids. At the same time, albumin synthesis within the hepatocytes is slowed. Thus, the redirection of hepatic secretory protein synthesis during infection is analogous to that seen after severe burns, which according to Cuthbertson (30) "requires a stimulatory factor from the injured site and the presence of a permissive level of cortical hormone. The increase in protein synthesis in the liver occurs in response to the formation there of messenger RNA."

HORMONAL RESPONSES

Many of the nonspecific generalized host metabolic responses to an infectious process are modulated by hormonal actions. Individual endocrine responses during infection are of variable duration and magnitude. However, the role of some hormones is not as prominent as after surgery or trauma (23).

Glucocorticoid Responses

Increased secretions of adrenal glucocorticosteroids begin shortly before the onset of fever (23). Despite limitations in available methodology for determining rates of cortisol secretion under non-steady-state conditions, available estimates suggest that cortisol production is increased during acute infection, but rarely to values greater than two to five times normal. In keeping with these estimates, plasma cortisol measured during morning hours in an uncomplicated systemic infection usually maintains concentrations equal to or only slightly above the normal peak morning values. During acute illness, however, cortisol production appears to lose its circadian nadir, with the result that normal morning values are maintained throughout the day (23). Although markedly elevated plasma cortisol concentrations are rare during most acute infections, the onset of septic shock or hepatic dysfunction may lead to steadily increasing plasma cortisol concentrations because it is degraded less rapidly. On the other hand, glucocorticoid production may cease if an infection is complicated by hemorrhagic infarction of the adrenal cortex.

The typical increase in glucocorticoid secretion is accompanied by smaller increases in the output of adrenal ketosteroids and pregnanetriol. These ACTH-mediated adrenocorticoid increases usually revert abruptly to normal with the onset of recovery. Should an infection become subacute or chronic, the excretion of adrenal glucocorticoids and ketosteroids generally falls to values somewhat below normal (23).

Responses Related to Salt and Water Metabolism

The extracellular electrolytes of the body are subjected to a complex variety of sequential changes (1, 2, 31). Increased renal losses of Na and Cl have been detected during a brief period prior to the onset of fever (2). Na, Cl, and K can also be lost directly through the sweat, diarrheal stools, or vomitus. The concentrations of Na in plasma may also be diminished because of expansion of the extracellular space and accumulation of Na within cells.

Increases in aldosterone secretion do not coincide in timing with those of cortisol (23), but become evident after fever has begun. An increased secretion of aldosterone then persists into early convalescence and abates gradually. With increased aldosterone secretion, both Na and Cl are avidly retained by the kidneys and may virtually disappear from the urine during periods of high fever. Secretion of antidiuretic hormone (which may become inappropriate) also leads to retention of water by the kidneys (31, 32). Extracellular water thus tends to accumulate in excess during fever; this excess is generally lost through a prominent period of diuresis which occurs early in the post-febrile period of convalescence (29).

Thyroid Hormone Responses

Thyroidal responses were poorly defined until they could be related to the stages of induced infection by means of sequential measurements in volunteers. This approach led to the realization that even naturally acquired types of infection were characterized by a biphasic pattern of thyroidal response (33–35).

Protein-bound iodine (PBI) values initially become depressed during the early phases of acute illness and then rebound above normal during convalescence. An accelerated rate of disappearance of thyroxine (T_4) from plasma (33–35), accompanied by increased fractional disappearance rates of both T_4 and triiodothyronine (T_3), and the absence of prompt pituitary-thyroidal compensation accounts for low PBI values during early infection. Enhanced secretion of T_4 takes place following the acute illness and persists for a time during convalescence. In patients with malaria, the fractional disappearance rate for T_4 is slowed, possibly due to impaired cellular deiodination within the liver (34).

A consensus has not been reached as to whether some of the early changes in peripheral tissue utilization of thyroid hormones might be influenced by alterations in the thyroid-binding proteins of plasma, or possibly by inhibitors of T_4 binding (35). Another question that remains open despite considerable study deals with the possibility that iodine, derived from the degradation of thyroid hormones, might play an important bactericidal role within phagocytic cells (35).

Carbohydrate-Regulating Hormones

Hormones that influence carbohydrate metabolism are intimately involved in the host response. During the early febrile phases of bacterial and viral illnesses, fasting plasma concentrations of glucose, insulin, glucagon, and even growth hormone tend to be slightly increased (6, 36). If carbohydrate tolerance is measured by means of an intravenous glucose load at that time, the insulin responses and glucose disappearance rates show changes approaching those characteristic of maturity onset diabetes. In addition, elevated fasting glucagon values decline appropriately after the glucose load, but growth hormone secretion seems to undergo an acute paradoxical stimulation (36). Few quantitative details are available concerning catecholamine responses during infection.

Basic biochemical studies suggest that these combined hormonal responses set in motion molecular mechanisms to release glucose from stored glycogen, i.e. activation of adenylate cyclase, formation of cyclic AMP, and a resetting of the hepatic glycogen synthetase and phosphorylase enzyme systems to favor glycogenolysis (37–39). In addition, the liver increases its rates of gluconeogenesis from available substrates (1, 40). In the face of the increased energy demands of acute infection or severe trauma, amino acid utilization serves the immediate need despite its potentially wasteful consequences. An increased flux of gluconeogenic amino acids such as alanine from muscle-to-plasma-to-liver supports this activity (1, 27, 28, 40).

Based upon a tendency toward increased hyperglycemia and glycosuria during infection, diabetic patients have been thought to require greater amounts of insulin. Long et al (41) studied the metabolism of glucose-^{14}C in nondiabetic patients after major injury or infection and observed a doubling of glucose pool size as well as increased rates of glucose turnover and oxidation. These results suggested that hepatic gluconeogenesis was sufficiently great so that hyperglycemia occurred despite an increased utilization of glucose by the body. In studies using tagged glucose in rabbits with pneumococcal sepsis, Guckian (42) also found a normal or increased rate of glycolysis in peripheral tissues. Although lactate values were increased, Guckian found no change in ATP or ADP of blood or tissues (42). In contrast, during the terminal stages of overwhelming infection or endotoxemia in newborns with sepsis, or in patients with severe liver damage such as may occur with yellow fever or viral hepatitis, the ability to synthesize glucose may be impaired with the result that glycogen reserves in liver and skeletal muscles become depleted and hypoglycemia results (7–9). A specific point of biochemical failure has not as yet been identified to account for an agonal breakdown in carbohydrate synthesis (42).

SPECIFIC METABOLIC RESPONSES DURING INFECTION

Although the metabolic responses to infection do not occur as isolated events, many specific changes can best be approached individually when investigating their interrelationships.

Specific Amino Acid Changes

Closely linked to the altered rates of synthesis or degradation of protein are the influences of infectious illness on amino acid metabolism. The concentration of free amino acids in plasma changes rapidly after the initiation of an infectious process, as evidenced by the declining values for total as well as for most individual amino acids (4, 15, 27, 28). Since a decline in amino acid concentrations generally begins during incubating illness, it cannot be explained by a diminished dietary intake of protein or by an increased loss of amino acids in the urine.

On the other hand, amino acid decrements in plasma seem to result from the increased flux of amino acids from plasma to liver, before and during the acute phase of symptomatic illness (4, 27). Depressed values in plasma result when increased rates of flux into the liver for most individual amino acids exceed the increase in their output from muscle. The increased flux of amino acids into the liver during infection appears to result from two simultaneous stimuli: one from LEM and the other from hormones such as glucagon which act via the adenylate cyclase enzyme on hepatic cell membranes. Although these two stimuli function independently, their combined effects provide greater quantities of plasma amino acids for hepatic utilization. Individual ones may be incorporated without modification into newly synthesized hepatic proteins, or they may be deaminated for the synthesis of glucose or ketones, or degraded via a specific individual metabolic pathway.

Patterns of change in individual plasma amino acids have been shown to vary depending upon the type of microorganism producing an infection. This variability can best be explained by different rates of utilization of individual amino acids. An illustration of this is found in the increased rate of tryptophan metabolism via the kynurenine-niacin pathway (21). This infection-related change in the metabolic utilization of an essential amino acid is shown both by heightened urinary excretion of various kynurenine metabolites during different infections (especially typhoid fever), as well as by an early increase in the activity of tryptophan oxygenase in the liver of laboratory animals with induced infection (21).

Unlike the other free plasma amino acids, concentrations of phenylalanine appear to increase in plasma rather than to decrease. The phenylalanine: tyrosine ratio (43) has been shown to increase markedly in a number of infections and other febrile conditions. This unique response cannot be explained by a diminished renal excretion of phenylalanine or by an appreciably diminished rate of its conversion to tyrosine within the liver. Further, the output of phenylalanine from muscle is not proportionately greater than that of other amino acids (28). Thus, accumulation of phenylalanine in plasma during febrile illnesses seems to be the combined result of its comparatively small rate of incorporation into newly synthesized protein within the liver, and its unsuitability for use as a substrate for glucose production or most other metabolic processes.

Although lowered values for most amino acids in plasma is typical of the early periods of many infections, increased values have sometimes been noted (42). A massive input of amino acids into plasma caused by the destruction or dysfunction of body cells could exceed the capacity for amino acid uptake, degradation, or excretion. Such a mechanism would account for increased concentrations in free amino acid values during severe or terminal infections.

Lipid Metabolism

Although fat depots can supply body energy needs during prolonged periods of total starvation, the mechanisms needed to permit this adaptive response seem to require a period of time for full development. As with many other metabolic substrates or products, a biphasic pattern of change is seen if plasma lipids are measured sequentially throughout the course of an infectious process (44). In the presence of a diminished dietary intake and an increased energy requirement for calories, fat stores are called upon to supply some of the body needs. Free fatty acids disappear from plasma more rapidly than normal during the early stages of infection, and the rates of triglyceride and cholesterol synthesis within the liver are accelerated (42, 44, 45). Despite increased hepatic production rates for lipid moieties, the metabolic needs of peripheral tissues may initially increase even faster. As a result, concentrations of most lipids in plasma decline during the early stages of a generalized uncomplicated infection. Such responses, however, are undoubtedly influenced by age, nutritional status, and by preexisting genetic capabilities of the host for handling various lipids. The availability of carbohydrates and amino acids as alternative sources of energy may further modify the rates of lipid metabolism during infection.

On the other hand, the disposal of lipids may become reduced because of progressive decreases in post-heparin lipolytic activity (46). If an infectious process begins to overwhelm the host, peripheral tissues appear to lose their ability to employ lipids effectively as a source of metabolic energy, and lipids then accumulate in plasma (3). Thus, the terminal hypothermia that may accompany the final phases of an overwhelmingly severe infection can be ascribed to the inadequacy of glucose production in combination with an inability of body tissues to utilize all potentially available sources of metabolizable energy.

Mineral Metabolism

Changes in the balance of Mg and inorganic phosphate (P_i) generally parallel the characteristic patterns of loss exhibited by N and K (1). Ca balance on the other hand does not exhibit clearly defined patterns of change during most infections.

Plasma concentrations of Ca and Mg may decline slightly during the febrile period of an infection, possibly as a result of dilutional effects of retained extracellular water. On the other hand, during periods of rising fever, plasma concentrations of P_i may fall (1). Such hypophosphatemia is most severe during gram-negative bacterial sepsis. It is possible that diminution of P_i concentrations in plasma and urine during fever may occur, in part as a manifestation of respiratory alkalosis due to the hyperventilation typically associated with rapidly rising body temperature.

The plasma concentrations of Fe and Zn are diminished during acute infections (5). Although the Fe-binding capacity of transferrin is generally unchanged during early infection, Fe per se may virtually disappear from the plasma, especially during severe pyrogenic bacterial infections. Hypoferremia cannot be ascribed to a loss of Fe from the body; rather, there is an accumulation of hemosiderin-Fe in hepatic storage depots. This redistribution of body Fe can lead to anemia if infection becomes chronic.

Movement of Fe and Zn out of plasma is stimulated by LEM released by phagocytic cells (5, 25, 26). Zn also appears to accumulate in the liver. The infection-related changes in plasma Zn resemble in many respects the changes in Fe although the magnitude of fall in Zn is proportionately smaller. The timing of changes in both the metals, however, is relatively parallel.

Although hypoferremia is typical of most infections, hepatitis may lead to hyperferremia. This paradoxical response may be ascribed to a failure of damaged hepatocytes to accumulate or retain intracellular Fe (5).

Changes in plasma Cu concentrations are also characteristic of acute and chronic infections. Unlike the decline of Fe and Zn, however, values for Cu generally increase. This increase occurs somewhat later during infection than the changes of Fe and Zn, progresses more slowly, and persists well into convalescence. Increased concentrations of Cu in plasma can be ascribed almost entirely to an increased production of ceruloplasmin by the liver and its gradual disappearance which lasts well into convalescence (5).

SUMMARY

A generalized infectious process leads to a broad and complex array of metabolic responses within the host. Certain of these responses have a direct causal relationship to the interactions of body cells with invading microorganisms or their products, or to specific host defensive mechanisms. Other host metabolic responses have less clearly defined roles although they involve many tissues and seem to be initiated and regulated by certain hormones and hormone-like endogenous mediators. These latter responses contribute to the maintenance of body homeostasis, the provision of metabolizable energy to meet increased body needs, and the synthesis of the acute phase reactant serum proteins.

Literature Cited

1. Beisel, W. R. 1972. *Am. J. Clin. Nutr.* 25:1254–60
2. Beisel, W. R., Sawyer, W. D., Ryll, E. D., Crozier, D. 1967. *Ann. Intern. Med.* 67:744–79
3. Gallin, J. I., Kaye, D., O'Leary, W. M. 1969. *N. Engl. J. Med.* 281:1081–86
4. Wannemacher, R. W. Jr. et al 1972. *J. Infec. Dis.* 126:77–86
5. Beisel, W. R., Pekarek, R. S., Wannemacher, R. W. Jr. 1974. In *Trace Element Metabolism in Animals - 2,* ed. W. G. Hoekstra, J. W. Suttie, H. Ganther, W. Mertz, Chap. 16, 217–40. Baltimore: Univ. Park Press
6. Rayfield, E. J., Curnow, R. T., George, D. T., Beisel, W. R 1973. *N. Engl. J. Med. 289*: 618–21
7. Yeung, C. Y. 1970. *J Pediat.* 77:812–17
8. LaNoue, K. F., Mason, A. D. Jr., Daniels, J. P. 1968. *Metabolism* 17: 606–11
9. McCallum, R. E., Berry, L. J. 1973. *Infec. Immunity* 7:642–54
10. Gordon, J. E., Scrimshaw, N. S. 1970. *Med. Clin. N. Am.* 54:1495–1508
11. Beisel, W. R. 1974. Nutrient wastage during infection. In *Proc. 9th Int. Congr. Nutr., Mexico, 1972.* Springfield, Ill.: Thomas. In press
12. Beisel, W. R., Goldman, R. F., Joy, R. J. T. 1968. *J. Appl. Physiol.* 24: 1–10
13. Young, V. R., Haverberg, L. N., Bilmazes, C., Munro, H. N. 1973. *Metabolism* 22:1429–36
14. Nessan, V. J., Geerken, R. C., Ulvilla, J. 1974. *J. Clin. Endocrinol. Metab.* 38: 652–54
15. Wannemacher, R. W. Jr., Dinterman, R. E., Pekarek, R. S., Beisel, W. R. 1974. *Fed. Proc.* 33:669 (Abstr.)
16. Klebanoff, S. J. 1971. *Ann. Rev. Med.* 22:39–62
17. Janoff, A. 1972. *Ann. Rev. Med.* 23: 177–90
18. Ginsberg, H. S. 1970. In *Infectious Agents and Host Reactions,* ed. S. Mudd, Chap. 20, 466–86. Philadelphia: Saunders
19. Mackaness, G. B. 1970. Ibid, Chap. 3, 61–75
20. Cline, M. J. 1965. *Physiol. Rev.* 45:674–720
21. Rapoport, M. I., Beisel, W. R. 1971. *Am J. Clin. Nutr.* 24:807–14
22. Powanda, M. C., Wannemacher, R. W. Jr., Cockerell, G. L. 1972. *Infec. Immunity* 6:266–71
23. Beisel, W. R., Rapoport, M. I. 1969. *N. Engl. J. Med.* 280:541–46, 569–604
24. Wood, W. B. Jr. 1970. See Ref. 18, Chap. 6, 146–62
25. Pekarek, R. et al 1974. *Life Sci.* 14:1765–76
26. Kampschmidt, R. F. 1974. *Ann. Okla. Acad. Sci.* 4:62–67
27. Wannemacher, R. W. Jr., Pekarek, R. S. 1973. *Fed. Proc.* 32:953 (Abstr.)
28. Wannemacher, R. W. Jr. et al 1974. *Am. J. Clin. Nutr.* 27:439 (Abstr.)
29. Earp, H. S. III, Pekarek, R. S. 1974. *Clin. Res.* 22:440 (Abstr.)
30. Cuthbertson, D. P. 1970. *J. Roy. Coll. Surg. Edinb.* 15:239–49
31. Bilbrey, G. L., Beisel, W. R. 1973. *Ann. Surg.* 177:112–20
32. Nyhan, W. L., Cooke, R. E. 1956. *Pediatrics* 18:604–13
33. Lutz, J. H., Gregerman, R. I., Spaulding, S. W., Hornick, R. B., Dawkins, A. T. Jr. 1972. *J. Clin. Endocrinol. Metab.* 35:230–49
34. Wartofsky, L., Martin, D., Earll, J. M. 1972. *J. Clin. Invest.* 51:2215–32
35. DeRubertis, F. R., Woeber, K. A. 1973. *J. Clin. Invest.* 52:78–87
36. Rayfield, E. J., George, D. T., Beisel, W. R. 1974. *J. Clin. Endocrinol. Metab.* 38:746–54
37. Tsung, P. K., Paretsky, D. 1968. *Acta Virol.* 12:49–53
38. Curnow, R. T., Rayfield, E. J. 1973. *Fed. Proc.* 32:954 (Abstr.)
39. Zenser, T. V., DeRubertis, F. R. 1973. *Fed. Proc.* 32:467 (Abstr.)
40. Felig, P. 1973. *Metabolism* 22:179–207
41. Long, C. L., Spencer, J. L., Kinney, J. M., Geiger, J. W. 1971. *J. Appl. Physiol.* 31:110–16
42. Guckian, J. C. 1973. *J. Infec. Dis.* 127:1–8
43. Newberne, P. M. 1966. *Fed. Proc.* 25:1701–10
44. Lees, R. S., Fiser, R. H. Jr., Beisel, W. R., Bartelloni, P. J. 1972. *Metabolism* 21:825–33
45. Fiser, R. H., Denniston, J. C., Beisel, W. R. 1972. *J. Infec. Dis.* 125:54–60
46. Kaufmann, R. L., Oland, D. D., Matson, C. F., Rowberg, A. H., Beisel, W. R. 1974. *Clin. Res.* 22:472 (Abstr.)

Copyright 1975. All rights reserved

X-LINKAGE IN MANIC-DEPRESSIVE ILLNESS

◆7127

Remi J. Cadoret, M.D. and George Winokur, M.D.
Department of Psychiatry, University of Iowa College of Medicine, Iowa City, Iowa 52242

Affective disorders comprise a large proportion of psychiatric illnesses, with morbid risk estimates in the general population running as high as 6–8% (1). These conditions primarily involve severe mood changes, and their characteristic course, as shown long ago by Kraepelin (2), is episodic with remission to the person's usual personality and psychological state between episodes. Mood changes take two directions: depressive, by far the most common, and manic, which is considerably rarer.

In recent years considerable evidence has separated affective disorders into two separate illnesses: bipolar and unipolar (3–5). In the bipolar condition, mania occurs sometime during the course of the affective illness; in the unipolar, only depressive episodes occur. Evidence supporting this differentiation is of two kinds: (*a*) clinical data, which show differences between unipolar and bipolar illnesses in such variables as length of episode, number of episodes, and age of onset of the condition (4, 6, 7), and (*b*) familial data, which show that bipolar families have higher rates of psychosis, (3, 8–11) especially manic episodes (4, 6). Most important to the present paper, bipolar families have shown in some series a *pattern of inheritance* of affective illness suggestive of X-linked dominance (12, 13).

Dominant X-linked transmission should manifest itself in a number of ways. One consequence is that more females than males should be affected; in the general population this is found in two large epidemiological studies (1, 14). Most series of bipolar patients do show an excess of ill females (4, 6, 12). Also, in the families of bipolar probands more females than males should be affected. In a comprehensive study of depressive illness, Winokur, Clayton & Reich found this to be true (15). If the sex ratio (living males: living females) in families with bipolar illness shows a preponderance of females, findings such as this would be questionable. In a small study, Brown et al (16) showed an excess of all female relatives over male relatives, but, in other studies, (4, 13) the number of female relatives equals the number of male relatives. Another consequence of X-linked dominant inheritance is that ill males should have equal numbers of affected brothers and sisters; ill females should

P, 24843 . ~~1634~~

have more ill sisters than brothers. This too was found by Winokur, Clayton & Reich (15). No father-to-son transmission of illness is a crucial consequence of an X-linked dominant condition. This is such an important point that we discuss it more completely below. The strongest point in favor of X-linkage would be the finding of linkage of the illness with a known genetic marker on the X chromosome.

The most direct evidence of X-chromosome involvement with bipolar illness has come from linkage studies. In situations where two genes on the same chromosome are close enough that crossing-over during meiosis is infrequent, it is possible to trace both genes through a number of generations because of that close linkage, and thus demonstrate that the two genes are on the same chromosome. Two objectively measured genetic traits have been found to be linked to bipolar illness. These markers, the Xg^a antigen and red-green color blindness (both protan and deutan types), are both located on the X chromosome (17, 18). Two groups of investigators have published studies involving these markers and all published data positively demonstrate linkage between one of these genetic markers and bipolar affective disorder. Results from these studies (17–20) appear in Table 1. By combining the data of studies 2 and 3, Mendlewicz et al (21) demonstrated linkage by estimating the recombination fraction θ (the relative frequency of "recombinants," or offspring demonstrating crossing-over to the number of nonrecombinants). A θ near zero implies very close linkage; one at 50%, no linkage. Mendlewicz et al reported a θ of $.11 \pm .08$ ($\theta \pm$ standard error) for protan color blindness and $.19 \pm .10$ for deutan color blindness. Both these values are significantly different from $\theta = 50\%$ and are consistent with genetic linkage. These same investigators (22) have computed θ values for linkage of Xg^a with manic-depressive illness and report $\theta = .23$ with a standard error of .10. Again, this is consistent with linkage.

Although Xg^a and the color blindness loci are on the X chromosome, they do not show linkage because the long distance between them allows for frequent crossing-over (23). The manic-depressive locus shows linkage with both Xg^a and the color blindness loci and thus is unique in bridging the gap between Xg^a and its linked conditions (ocular albinism and ichthyosis) and deutan color blindness with its linked conditions (hemophilia, protan color blindness, g-6-pd). The genetic map distance between g-6-pd and Xg^a appears to be at least 40–50 map units (21) since linkage is not detectable, which suggests that the manic-depressive locus could well fit in the intervening space.[1]

Inspection of pedigrees from studies in Table 1 suggests a dominant-type transmission, although Reich, Clayton & Winokur (17) argue cogently for a modified sex-linked dominance involving either (*a*) diminished penetrance of one gene or (*b*) more than one dominant gene necessary for manifestation of the illness. At the present time the question of penetrance or a second gene is unresolved (13).

---

[1] Recent weak evidence cited by McKusick (Joseph C. Wilson lecture, University of Rochester, Nov. 1, 1973) suggests that the Xg^a locus may be on the short arm of the X chromosome. If this is true, it is unlikely that manic-depressive illness could be linked both to Xg^a and color blindness loci. We must await more definitive data on the location of Xg^a.

Table 1 Linkage studies in bipolar affective disorder

Study	Genetic Marker	Results
1. Winokur & Tanna (18)	Xg^a	Report 3 pedigrees informative for linkage. In these, 10 children were compatible with X-linked transmission, one was incompatible. This difference is significant (2 of the 3 families suggest specific linkage with Xg^a blood system. This was not significant, however).
2. Reich, Clayton & Winokur (17)	red-green color blindness	Report 2 pedigrees, one with protan, the other with deutan color blindness. In each pedigree significant evidence of linkage with color blindness was found. Pedigrees suggest dominant inheritance.
3. Fieve et al (19)	red-green color blindness	Report 8^a pedigrees—4 with protan and 4 with deutan color blindness. Total of protan and deutan families gave significant evidence for linkage. Pedigrees fit a dominant inheritance.
4. Mendlewicz, Fleiss & Fieve (20)	Xg^a	Report 12^a pedigrees for Xg^a. Significant linkage reported.

[a] [illegible]

Table 2 Father-to-son "transmission" of bipolar affective disorder

Series Compatible with X-linked Transmission	Results
1. Winokur, Clayton & Reich (15)	In 89 families *no* ill father-son pairs.
2. Taylor & Abrams (24)	In 55 families *no* ill father-son pairs.
Series Not Compatible with X-linked Transmission as the *Sole* Transmission Possibility	
1. Perris (25)	In series of 138 bipolar probands, 13 ill father-son pairs.
2. Dunner, Gershon & Goodwin (26)	In 23 bipolar male patients, 4 had affectively ill fathers.
3. Green et al (27)	In 35 bipolar families, 4 ill father-son pairs (no illness on maternal side).
4. Fieve et al (19)	In 120 bipolar probands, 13 ill father-son pairs (9 pairs no evidence of illness on maternal side).
5. Helzer & Winokur (28)	Of 30 male manics, 8 affectively ill mothers (41%), and 1 affectively ill father (5%); in another ill father-ill son pair, considerable affective disorder existed on the maternal side of the family.
6. Von Greiff, McHugh & Stokes (29)	In 16 male probands, 4 ill father-son pairs with *no* illness on maternal side; 3 ill father-son pairs did have illness on maternal side.

which is inherited in some families in an X-linked fashion and in others as an autosomal type (30). The importance of the present linkage studies lies in the fact that a genetic etiology in some families has been demonstrated for a major psychiatric illness. This has some limited practical application at present in genetic counseling (31).

Literature Cited

1. Helgason, R. 1964. *Acta Psychiat. Scand.* 40 (Suppl. 173):258
2. Kraepelin, E. 1901. *Psychiatrie: ein Lehrbuch für Studierende und Ärtze.* Leipzig: Barth. 9th ed.
3. Leonhard, K. 1966. *Aufteilung der Endogenen Psychosen.* Berlin: Akad-Verlag
4. Perris, C. 1966. *Acta Psychiat. Scand. Suppl.* 194:420
5. Winokur, G., Clayton, P. 1967. *Recent Advances in Biological Psychiatry,* Vol. 9. New York: Plenum
6. Angst, J. 1966. *Monographien aus dem Gesamtgebiete der Neurologie und Psychiatrie,* Vol. 112. Berlin: Springer
7. Cadoret, R., Winokur G. 1972. *Int. J. Ment. Health.* 1:159–75
8. Asano, N. 1967. In H. Mitsuda, Ed. *Clinical Genetics in Psychiatry.* Tokyo: Igaku Shoin
9. Dorzab, J., Baker, M., Cadoret, R., Winokur, G. 1971. *Am. J. Psychiat.* 127:1128–33
10. Kinkelin, M. 1954. *Schweiz. Arch. Neurol. Psychiat.* 73:100–146
11. Neele, E. 1949. *Die phasischen Psychosen nach ihrem Erscheinungs und Erbbild.* Leipzig: Barth
12. Cadoret, R., Winokur, G., Clayton, P. 1970. *Brit. J. Psychiat.* 116:625–35
13. Winokur, G., Reich, T. 1970. *Compr. Psychiat.* 11:93–99
14. Fremming, K. H. 1951. *The Expectation of Mental Infirmity in a Sample of the Danish Population.* London: Cassell

15. Winokur, G., Clayton, P., Reich, T. 1970. *Manic Depressive Disease.* St. Louis: Mosby
16. Brown, R., Elston, R. C., Pollitzer, W. S., Prange, A., Wilson, I. 1973. *Biol. Psychiat.* 6:307–9
17. Reich, T., Clayton, P., Winokur, G. 1969. *Am. J. Psychiat.* 125:1358–69
18. Winokur, G., Tanna, V: L. 1969. *Dis. Nerv. Syst.* 30:89–93
19. Fieve, R., Mendlewicz, J., Rainer, J., Fleiss, J. 1974. Linkage studies in affective disorders, II. Color blindness and manic-depressive illness. In *Genetics and Psychopathology,* ed. R. Fieve, D. Rosenthal, H. Brill. Baltimore: Johns Hopkins Univ.
20. Mendlewicz, J., Fleiss, J., Fieve, R. 1974. Linkage studies in affective disorders, I. Xg blood group and manic-depressive illness. In *Genetics and Psychopathology,* ed. R. Fieve, D. Rosenthal, H. Brill. Baltimore: Johns Hopkins Univ.
21. Mendlewicz, J., Fleiss, J. L., Fieve, R. R. 1972. *J. Am. Med. Assoc.* 222:1624–27
22. Fieve, R. R., Mendlewicz, J., Fleiss, J. L. 1973. *Am. J. Psychiat.* 130:1355–59
23. Race, R. R., Sanger R. 1968. *Blood Groups in Man.* Philadelphia: Davis. 5th ed.
24. Taylor, M., Abrams, R. 1973. *Arch. Gen. Psychiat.* 28:656–58
25. Perris, C. 1971. *Brit. J. Psychiat.* 118:207–10
26. Dunner, D., Gershon, E., Goodwin, F. K. 1969. *Heritable Factors in the Severity of Affective Illness.* Presented at 123rd Ann. Meet. APA, San Francisco
27. Green, R., Goetze, V., Whybrow, P., Jackson, R. 1973. *J. Am. Med. Assoc.* 223:1289
28. Helzer, J., Winokur, G. W. 1974. *Arch. Gen. Psychiat.* 31:73–77
29. Von Greiff, H., McHugh, P. R., Stokes, P. 1973. *The Familial History in Sixteen Males with Bipolar Manic-Depressive Disorder.* Presented at 63rd Ann. Meet. Am. Psychopathol. Assoc., New York
30. Stern, C. 1960. *Principles of Human Genetics.* San Francisco:Freeman. 2nd ed.
31. Cadoret, R. 1973. *The Genetics of Affective and Genetic Counseling.* Presented at Workshop on Genetic Counseling, Univ. Indiana, Indianapolis, Indiana
32. Mendlewicz, J., Fleiss, J. 1974. Linkage studies with X-chromosome markers in bipolar (manic-depressive) and unipolar (depressive) illness. *Recent Advan. Biol. Psychiat.* In press

Copyright 1975. All rights reserved

SKIN LIPIDS AND ACNE[1]

❖7128

John S. Strauss, M.D., Peter E. Pochi, M.D., and Donald T. Downing, Ph.D.[2]

Department of Dermatology, Boston University School of Medicine and Evans Memorial Department of Clinical Research, University Hospital, Boston, Massachusetts 02118

INTRODUCTION

The presence of a lipid mantle is characteristic of integumentary surfaces. In species other than man, it serves a number of biologic functions. As an example, the preen gland, the only cutaneous gland in aquatic birds, is essential for protection against the macerating effect of water on feathers. In these birds, the lipid is also essential for protection against heat loss. Modified sebaceous glands also serve in various mammals as a source of pheromones involved in sexual attraction and territorial marking (1). While it is conceivable that the presence of skin lipids is in some way necessary for the functional maintenance of the skin, no significant physiological importance has been demonstrated for skin lipids in man. It is thus ironic and unfortunate that the glands, which enlarge at puberty and constitute the major source of the skin lipid, should have no protective function and yet be an integral factor in the pathogenesis of acne. It is this relationship that will be examined in this review. However, it is first necessary to present information concerning the source and composition of the cutaneous lipids in order to understand the role of sebaceous gland lipids in acne.

THE SOURCE OF SKIN SURFACE LIPID

The lipids recovered from the skin surface are derived from two sources, the sebaceous glands and the epidermis. The sebaceous glands, which are hormone-dependent for their development and functioning, form their secretory product, sebum, by a holocrine process in which the individual cells accumulate lipid as they mature in a centripetal direction within the glandular acini. The lipid-laden cells rupture, discharging their contents to the skin surface via the excretory canal system of the pilosebaceous unit. Sebaceous gland morphology varies regionally and with age. The number and size of glands are greatest on the scalp, face, and upper trunk and are smallest on the acral areas. They are well-developed at birth, but shortly thereafter decrease in size and remain small until puberty, when they enlarge chiefly as a result of androgenic stimulation (2). Sebaceous gland activity begins to decline after menopause in women and after the seventh decade in men (2).

---

[1] Supported in part by National Institutes of Health Grants AM 07388 and AM 07084.

[2] Dr. Downing is a recipient of National Institutes of Health Career Development Award AM 42402.

It is important, therefore, to be aware of these age- and area-dependent factors in sebaceous gland development, as the surface lipid is a mixture derived both from the glands and epidermal cells. The surface lipid from gland-rich areas naturally contains a higher proportion of sebaceous lipid, whereas from gland-deficient areas such as the arms and legs there is a greater proportion of epidermally derived lipid. That the epidermal cells contribute to the lipid film is not surprising. They are pluripotential and can form lipid as well as proteins during the process of keratinization. Under normal circumstances, lipid droplets are not visible in epidermal cells, but in certain pathological conditions, such as psoriasis, lipid may be visible at the ultrastructural level. The epidermal lipid is probably delivered to the surface at a constant rate as the horny cells mature, and when sebaceous gland development is minimal, as in the prepuberal child, it constitutes the major source of the surface lipid.

Pure sebum can be obtained by dissection of the sebaceous glands (3); however, dissecting out the glands may be difficult and only very small samples are obtainable for study. To date, it has not been possible to collect sebum by cannulation of the pilosebaceous follicle. Therefore, surface lipid collected from sebaceous-rich areas such as the scalp and face represents the purest "sebum" that can be obtained in reasonable amounts. In adults, the amount of surface lipid on the forehead, for example, normally varies from 150–300 μg of lipid per square centimeter, of which the epidermal contribution is only 5–10 μg, or 3–6% (4).

THE COMPOSITION OF SKIN SURFACE LIPID

Skin lipids are a complex mixture of triglycerides, fatty acids, wax esters, squalene, cholesterol, and cholesterol esters. Moreover, some major constituents are not found elsewhere in the body; thus, information concerning biosynthetic pathways in other tissues does not necessarily apply to the skin. The fact that the skin has two sources of lipid (glands and epidermis) likewise makes experiments involving whole skin difficult to interpret unless isolated glands or epidermis are used. Epidermal tissue, such as from guinea pig ear skin, has been used to elucidate the biosynthetic pathways (5), but because of species differences the data cannot properly be extrapolated to human skin. While studies of the preputial gland in rats indicate that sterol synthesis in this sebaceous gland homologue might proceed via the Kandutsch-Russell pathway (6), there is no evidence for this in man, where sterol synthesis is minimal and there is accumulation of the sterol precursor, squalene.

Nonetheless, by study of the composition of the surface film, considerable insight is afforded concerning the origin and biosynthesis of human skin lipids (7). In order of decreasing concentration, surface lipid is composed of triglycerides and their hydrolysis products (diglycerides, monoglycerides, and free fatty acids), wax esters, squalene, sterol esters, and free sterols. The average composition of the surface film in adults is as follows: triglycerides 41.0%, wax esters 25.0%, free fatty acids 16.4%, squalene 12.0%, diglycerides 2.2%, cholesterol esters 2.1%, and cholesterol 1.4% (8). The presence of wax esters and squalene and of free fatty acids in a relatively high percentage distinguishes sebum from lipids in other mammalian tissues.

Insofar as the source of the individual classes of lipid is concerned, it has long been recognized that the cholesterol and squalene content varies depending upon the amount of surface lipid (9). Since the epidermal contribution to the surface lipid film is presumably relatively constant, an increase in the amount of surface lipid results from an increased production of sebum. When the amount of surface lipid is low, the cholesterol concentration is high and the squalene concentration low (9). Therefore, cholesterol is predominantly of epidermal origin, whereas squalene is derived from the sebaceous glands. There is additional evidence to support these conclusions. First, chemical analysis of sebaceous glands isolated by microdissection reveals the presence of only squalene, wax esters, and triglycerides (3). Second, the relative composition of the surface lipid, when examined in individuals from birth through adolescence, shows striking changes in composition with age (10). Cholesterol values as high as 30% are observed prior to puberty. The percentage of wax esters, which are sebaceous gland lipids, varies from 4–10% during childhood, increasing to 25–30% after puberty. The squalene concentration is more variable, but in general also increases at puberty.

Of the lipid constituents, most attention has been focused over the years on the fatty acids of sebum because of their putative role in the pathogenesis of acne (see below). The fatty acids of sebum are unique in that there are significant amounts of both branched-chain and odd carbon number acids. Since it has been shown that vernix caseosa, the fetal skin surface lipid, contains a much higher proportion of branched-chain acids than does adult surface lipid (78% vs 12%), it is unlikely that these acids are of bacterial origin, as vernix is sterile (11). The exact biosynthetic pathways involved in the genesis of the branched-chain fatty acids are not known, but these compounds are probably primarily of epidermal origin (7) and may thus serve to identify further the relative contribution of sebaceous glands and epidermis to the surface lipid mantle.

The fatty acids, both free and esterified, contain approximately equal proportions of saturated and unsaturated compounds. The double bonds of the unsaturated acids are predominantly in the Δ6 or Δ8 positions, which are not found in fatty acids from other animal tissues (7). These fatty acids are of sebaceous gland origin. A small proportion of fatty acids with Δ9 unsaturation is present; these are probably derived from epidermal cells. As expected, then, nearly all the unsaturated fatty acids in wax esters have Δ6 or Δ8 unsaturation. In contrast, the unsaturated fatty acids in cholesterol esters contain approximately 70% Δ9 and 30% Δ6 or Δ8 double bonds, indicating that approximately two thirds of the sterol esters are of epidermal origin but that there is some sterol ester synthesis in the sebaceous glands (7). Hydroxy

UNIVERSITY COLLEGE OF MEDICINE
LIBRARY
C. U.

acids, present in vernix, have been isolated from the outer horny layer and therefore are of epidermal origin (12).

Free fatty acids are not detectable in the lipids from isolated sebaceous gland tissue (3), but are found in varying percentages on the skin surface; the amount increases with the length of time that the lipid is allowed to accumulate on the skin (13). The following evidence indicates that the only source of these free fatty acids is the triglycerides. First, and most important, the percentages of triglycerides and free fatty acids in surface lipid samples are inversely proportional (8, 13, 14). In addition, small amounts of the intermediary products of triglyceride hydrolysis (diglycerides and monoglycerides) are found in the surface lipid (8). Other possible sources for the free fatty acids of the surface lipids are the wax esters and cholesterol esters. Hydrolysis of wax esters is unlikely since free fatty alcohols are not detectable in skin lipids, and the free cholesterol present would account for only a small proportion of the free fatty acids. Since the free fatty acids are derived from triglycerides and since the extent of hydrolysis of the triglycerides varies from person to person, it is useful to group the triglycerides and their hydrolysis products—diglycerides, monoglycerides, and free fatty acids—as "glyceride fatty acids" (14). The approximate composition of sebum is then calculated to be as follows: glyceride fatty acids 57.5%, wax esters 26.0%, squalene 12.0%, cholesterol esters 3.0%, and cholesterol 1.5% (4). By comparison, surface epidermal lipid is comprised of 65% glyceride fatty acids, 15% cholesterol esters, and 20% cholesterol (7). A small amount of cholesterol may be synthesized in sebaceous glands but no wax esters or squalene are found in epidermal lipid (4).

Over the past few years, a large body of evidence, which has been summarized recently (15, 16), has shown that the triglyceride hydrolysis is the result of microbial action. Organisms known to reside in the follicular reservoir into which sebum is secreted have the capability of hydrolyzing sebum triglycerides in vitro. These organisms, all normal residents, include the abundant anaerobe *Corynebacterium acnes,* staphylococci, micrococci, and pityrosporum species. In vivo studies, however, suggest strongly that the corynebacteria are the only significant source of intrafollicular hydrolysis (17–20). Their numbers are reduced by the administration of broad-spectrum antibiotics, with a resultant decrease in the formation of free fatty acids (19).

THE RELATION BETWEEN SEBUM AND ACNE

There is increasing evidence that the primary alteration in acne is an abnormal epithelial development in the pilosebaceous follicles (21, 22). The sebaceous glands themselves show no pathologic change; strictly speaking, therefore, the frequent assertion that acne is a sebaceous gland disorder is incorrect. Nonetheless, there are several compelling observations indicating the important involvement of the sebaceous glands in the acne process:

(*a*) From a hormonal standpoint, the development of the sebaceous gland and the occurrence of acne are both dependent upon the presence of androgen (2).

(*b*) Sebaceous gland secretion in acne is generally increased, particularly with greater severity of the disease (2, 23). However, since values for sebum excretion in acne patients and normal individuals overlap, it is not possible to distinguish the acne patient by the measurement of sebum production (23).

(*c*) Sebum, and its free fatty acids, are comedogenic; when applied to the inner surface of the rabbit ear they produce comedones, the keratinized impactions of the follicle characteristic of acne (24–27).

(*d*) Sebum is an irritant when injected intradermally (28) or applied topically under experimental conditions of occlusion (29).

The free fatty acids, particularly those with chain lengths of C_{10} and C_{12}, produce the most inflammation on topical application (29). The free fatty acids also produce inflammation when they are injected intradermally (28). It is likely that the inflammatory changes in acne result from the escape of sebum and its free fatty acids into the dermis. Certain therapeutic modalities of proven benefit in acne either decrease sebum production or decrease triglyceride lipolysis. Agents decreasing the production of sebum include estrogens (30–32), superficial X-irradiation (33), and an inhibitor of lipid synthesis (34). While broad-spectrum antibiotics do not affect sebum production, they decrease the hydrolysis of triglycerides to free fatty acids (35–37). Since the fatty acids are possibly comedogenic and have the capacity to provoke an inflammatory response, a decrease in their formation would be expected to be helpful therapeutically.

There are several reports describing the composition of sebum in acne (38–45). Unfortunately, these studies do not allow ready comparison with each other, since most of the investigators have not analyzed all of the constituents of sebum. Also, it is difficult to explain findings such as the report by one investigator in acne that the free fatty acids are decreased in men but are elevated in women (41). There is general agreement, however, that the free fatty acid content of sebum is not elevated in acne, although this does not rule against their involvment in the pathogenesis of this disease since, as noted previously, sebum secretion is increased in acne. As a result, the absolute amount of free fatty acids is higher than normal. In only three studies has squalene been measured (43–45), and in each case it was elevated, a probable reflection of an increase in sebum production. A minor fatty acid, identified as octadeca-5, 8-dienoic acid has been found to be present in higher amounts in the sebum of acne patients (46, 47). However, this component usually accounts for less than 1% of the total fatty acids in surface samples, and its significance is as yet unknown.

On the basis of current knowledge, therefore, it may be concluded that the alterations of sebum in acne appear to be largely quantitative rather than qualitative. However, the analyses reported to date have in every case dealt with the examination of surface lipid, and it is entirely possible that different relationships may obtain within the follicle, the site of the pathologic events in acne. Since only a small proportion of the pilosebaceous units are involved in the acne process at any time, major deviations of lipid composition might occur at the site of the developing lesion without influencing the composition of the surface lipid.

Literature Cited

1. Mykytowicz, R., Goodrich, B. S. 1974. *J. Invest. Dermatol.* 62:124–31
2. Pochi, P. E., Strauss, J. S. 1974. *J. Invest. Dermatol.* 62:191–201
3. Kellum, R. E. 1967. *Arch. Dermatol.* 95:218–20
4. Greene, R. S., Downing, D. T., Pochi, P. E., Strauss, J. S. 1970. *J. Invest. Dermatol.* 54:240–47
5. Wheatley, V. R. 1974. *J. Invest. Dermatol.* 62:245–56
6. Wilson, J. D. 1963. *Advances in Biology of Skin,* ed. W. Montagna, R. A. Ellis, A. F. Silver, 4:Chap. 10, 148–66. Oxford: Pergamon
7. Downing, D. T., Strauss, J. S. 1974. *J. Invest. Dermatol.* 62:228–44
8. Downing, D. T., Strauss, J. S., Pochi, P. E. 1969. *J. Invest. Dermatol.* 53:322–27
9. Boughton, B., MacKenna, R. M. B., Wheatley, V. R., Wormall, A. 1957. *Biochem. J.* 66:32–38
10. Ramasastry, R., Downing, D. T., Pochi, P. E., Strauss, J. S. 1970. *J. Invest. Dermatol.* 54:139–44
11. Downing, D. T. 1963. *Aust. J. Chem.* 16:679–82
12. Ansari, M. N. A., Fu, H. C., Nicolaides, N. 1970. *Lipids* 5:257–82
13. Nicolaides, N., Wells, G. C. 1957. *J. Invest. Dermatol.* 29:423–33
14. Lewis, C. A., Hayward, B. 1971. *Modern Trends in Dermatology,* ed. P. Barrie, 4:89–121. London: Butterworth
15. Shalita, A. R. 1974. *J. Invest. Dermatol.* 62:332–35
16. Strauss, J. S., Pochi, P. E., Downing, D. T. 1974. *J. Invest. Dermatol.* 62:321–25
17. Marples, R. R., Kligman, A. M., Lantis, L. R., Downing, D. T. 1970. *J. Invest. Dermatol.* 55:173–78
18. Marples, R. R., Downing, D. T., Kligman, A. M. 1971. *J. Invest. Dermatol.* 56:127–31
19. Marples, R. R., Downing, D. T., Kligman, A. M. 1972. *J. Invest. Dermatol.* 58:155–59
20. Marples, R. R. 1974. *J. Invest. Dermatol.* 62:326–31
21. Knutson, D. D. 1974. *J. Invest. Dermatol.* 62:288–307
22. Plewig, G. 1974. *J. Invest. Dermatol.* 62:308–15
23. Pochi, P. E., Strauss, J. S. 1964. *J. Invest. Dermatol.* 43:383–88
24. Kligman, A. M., Katz, A. G. 1968. *Arch. Dermatol.* 98:53–57
25. Lorincz, A. L., Krizek, H., Brown, S. 1968. *13th Congr. Int. Dermatol. Munich,* 1016–17. Berlin: Springer
26. Kligman, A. M., Wheatley, V. R., Mills, O. 1970. *Arch. Dermatol.* 102: 267–75
27. Kanaar, P. 1971. *Dermatologica* 142: 14–22
28. Strauss, J. S., Pochi, P. E. 1965. *Arch. Dermatol.* 92:443–56
29. Ray, T., Kellum, R. E. 1971. *J. Invest. Dermatol.* 57:6–9
30. Jarrett, A. 1955. *Brit. J. Dermatol.* 67:166–79
31. Strauss, J. S., Pochi, P. E. 1964. *J. Am. Med. Assoc.* 190:815–19
32. Pochi, P. E., Strauss, J. S. 1966. *J. Invest. Dermatol.* 47:582–85
33. Strauss, J. S., Kligman, A. M. 1959. *J. Invest. Dermatol.* 33:347–56
34. Strauss, J. S., Pochi, P. E., Whitman, E. N. 1967. *J. Invest. Dermatol.* 48:492–93
35. Freinkel, R. K., Strauss, J. S., Shin, Y. Y., Pochi, P. E. 1965. *N. Engl. J. Med.* 273:850–54
36. Strauss, J. S., Pochi, P. E., 1966. *J. Invest. Dermatol.* 47:577–81
37. Cunliffe, W. J., Cotterill, J. A., Williamson, B. 1972. *Brit. J. Dermatol.* 87: 37–41
38. Powell, E. W., Beveridge, G. W. 1970. *Brit. J. Dermatol.* 82:243–49
39. Šalamon, T., Bobarevié, B., Lazovíc, O., Pujić, Z., Hadžimusić, M. 1970. *Acta Dermato. Venereol.* 50:176–78
40. Bonelli, M., Cebrelli, G., Caputo, R., Piccinini, S. 1971. *G. Ital. Dermatol.* 46–112:120–25
41. Kanaar, P. 1971. *Dermatologica* 143: 121–29
42. Vecova, N., Picin, D. 1971. *Dermatol. Venerol. Sofia* 10:14–19
43. Cotterill, J. A., Cunliffe, W. J., Williamson, B., Bulusu, L. 1972. *Brit. Med. J.* 3:444–46
44. Gloor, M., Graumann, U., Kionke, M., Wiegand, I., Friederich, H. C. 1972. *Arch. Dermatol. Forsch.* 242: 316–22
45. Tronnier, H., Brunn, G. 1972. *Berufsdermatosen* 20:79–88
46. Kellum, R. E., Strangfeld, K. 1972. *J. Invest. Dermatol.* 58:315–18
47. Krakow, R., Downing, D. T., Strauss, J. S., Pochi, P. E. 1973. *J. Invest. Dermatol.* 61:286–89

Copyright 1975. All rights reserved

VON WILLEBRAND'S DISEASE

❖7129

Lars Holmberg, M.D. and Inga Marie Nilsson, M.D.[1]
Coagulation Laboratory, University of Lund, Allmänna Sjukhuset, Malmö, Sweden

INTRODUCTION

An abnormal bleeding tendency is often due to an inherited defect of hemostasis. One of the commonest of such defects is von Willebrand's disease (VWD), which is multifaceted but nevertheless has a typical clinical picture. Investigations of the cause of the condition have resulted in good prospects for controlling formerly life-threatening bleeding. Research in VWD also has contributed considerably to our knowledge of normal hemostasis.

HISTORY

In 1926 Erik Adolf von Willebrand (1) described an inheritable bleeding disease on the islands of Åland in the Gulf of Bothnia (Finland). Von Willebrand's description included bleeding from the gums and after tooth extraction, bleeding from the female genital tract, and bleeding from trivial wounds, but he stressed that joint bleedings, common in hemophilia, were relatively rare.

Von Willebrand distinguished the disease from anaphylactoid purpura and thrombocytopenic purpura as well as from the hereditary hemorrhagic thrombasthenia described by Glanzmann. He concluded that the condition was a previously unknown form of hemophilia which affected both sexes, and he called the disease hereditary pseudohemophilia. A prolonged bleeding time was its most important characteristic.

In 1928 the clinical picture was reported independently by four Americans, including Minot (2). In the 1930s Jürgens reinvestigated the original families on Åland and concluded that the platelets, as in Glanzmann's thrombasthenia, were defective (3). Other authors were, however, unable to demonstrate either a morphologic platelet defect or abnormal clot retraction in cases which in all respects resembled the disease described by von Willebrand (4, 5). They sought the cause of the disease in the morphology and function of the capillaries (4).

In the 1950s methods were developed for measuring the various hemophilia factors. In 1953 Alexander & Goldstein (6) described two cases with a combined

---

[1] Supported by a grant from the Swedish Medical Research Council (B75-19X-87-11A).

defect in hemostasis, namely, a prolonged bleeding time and decreased antihemophilic factor A (AHF) activity. Similar cases were later observed by a number of other authors; Schulman et al (7) called the disease vascular hemophilia, while Singer & Ramot (8) used the term, pseudohemophilia B.

In 1957 Nilsson et al (9) found the AHF activity to be reduced in 15 cases of VWD from the original families on Åland. This finding clearly showed that the bleeding disease observed in various quarters of the world was identical to that described by von Willebrand. The same group (10) found that the platelets were normal in respect to platelet factors 1, 3, and 4. The then novel AHF concentrate, plasma fraction I-0 (11), prepared from Cohn's fraction I, corrected not only the AHF deficiency, but also the bleeding time and the capillary bleedings.

PATHOGENESIS

Nilsson et al (12) later showed that patients with VWD lack a plasma factor occurring not only in normal plasma, but also in hemophilia plasma. This so-called v. Willebrand's factor (VWF) is not present in pure fibrinogen fractions or in platelets. The VWF not only corrects the prolonged bleeding time in patients with VWD, but also apparently stimulates the synthesis of AHF. The plasma fraction I-0, prepared from hemophilia plasma and thus devoid of AHF activity, increases the AHF activity when given to patients with VWD (12). Similar results have been reported by others (13, 14). The VWF appears to be closely related to AHF and occurs in AHF-rich plasma concentrates (12, 15).

It was at first not known how a plasma factor could affect primary hemostasis and shorten the bleeding time. The basis of primary hemostasis is the formation of a platelet plug. Borchgrevink (16) found decreased platelet adhesiveness in vivo in patients with VWD by comparing the platelet count in venous blood and in blood from a capillary lesion. A reliable in vitro test for measuring the interaction between platelets and a glass surface was devised by Hellem (17).

Hellem's original method, which used citrated blood, did not distinguish normal persons from those with VWD (18). Salzman (19) used native blood and a quicker flow and then could demonstrate a decreased platelet adhesiveness to glass in VWD. Modifications of Salzman's method, all with a rapid blood flow, have also been diagnostically useful (20, 21), although they are technically difficult and subject to methodological errors. Common to all methods of measuring platelet adhesiveness that can distinguish between persons with and without VWD is the use of a rapid flow of blood through a column of glass beads.

The reduced platelet adhesiveness in VWD, like the bleeding time, can be normalized by infusion of fresh platelet-poor plasma (19). Such normalization can be demonstrated in vitro by mixing the patient's blood with normal plasma, hemophilia plasma, or cryoprecipitate (22). The factor causing such normalization is missing in VWD plasma.

Further knowledge about the pathogenesis of VWD has been obtained in recent years since the advent of highly purified AHF. The AHF, as isolated from plasma, is a very large macroglobulin. Monospecific precipitating rabbit antisera against this globulin have been prepared. Patients with hemophilia A have a cross-reacting but inactive protein in the plasma (23, 24), while the majority of patients with VWD have a low content of this so-called AHF-related antigen (23, 25). The chromatographic pattern of normal (and hemophilia) cryoprecipitate on Sepharose 6-B according to v. Mourik & Mochtar (26) shows a typical protein peak corresponding to the AHF-related plasma protein in the void volume. No such peak is seen when the starting material consists of cryoprecipitate from a patient with severe VWD. The content of AHF-related protein is thus often really decreased in VWD (27; Holmberg & Nilsson, unpublished).

Bouma et al (27) showed that the AHF-related high molecular weight protein is responsible for enabling normal and hemophilic plasma to normalize the reduced platelet adhesiveness in VWD. High molecular weight AHF thus possesses not only AHF activity, but also VWF activity, and can then be called AHF-VWF protein.

What is the structure of this protein? Does it consist of two components, one possessing the AHF activity and the other the VWF activity, or does it consist of a single but bifunctional molecule (27)? Much evidence suggests that the protein consists of identical or at any rate equally large protein chains, possibly connected by covalent bonds to a large molecule and by noncovalent bonds to very large aggregates, which dissociate in various salt solutions (28–30). Other investigations, however, suggest that the protein complex consists of immunologically and biologically separate components (31–34).

Administration of AHF concentrate to a patient with severe VWD has, as pointed out above, a number of effects (Figure 1). It will shorten the bleeding time, and an increase in AHF activity out of proportion to the activity of the concentrate will occur. But this is not accompanied by any increase in the amount of antigenic

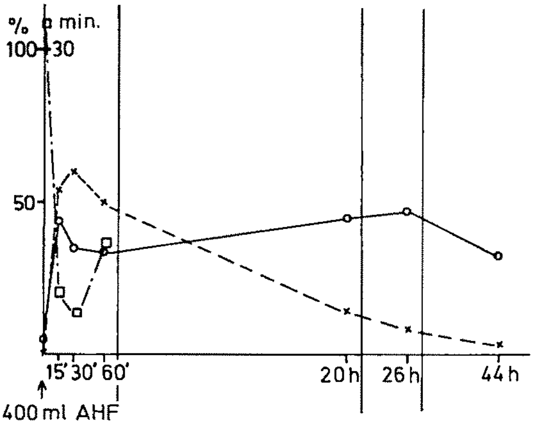


Figure 1 Infusion of AHF concentrate (fraction I–0) to a patient with severe v. Willebrand's disease. o—o AHF activity, x---x AHF-related antigen (AHF-VWF antigen), □--□ Duke's bleeding time.

protein, which instead decreases successively (35, 36). AHF activity may thus occur in some substance, which does not react to the antiserum with any demonstrable precipitation. This AHF activity may perhaps occur in a plasma fraction of lower molecular weight (37).

Judging from available experimental data, one might explain the defect in VWD as follows. The building stones of the high molecular weight AHF, i.e. protein chains with a molecular weight of 180,000–240,000 (28, 29, 38), are coded from an X-chromosomal AHF locus. The chains contain the sequences necessary for both AHF activity and VWF activity in the final molecule. But if the molecule is to have VWF activity, it must have a very special quarternary structure. The VWF activity seems to be labile, more labile than the ability of the protein to react with antiserum. Autosomal genes are responsible for polymerization of the chains to a large molecule with the special configuration.

In the most common form of VWD with a low content of AHF antigen (25), the defect may consist of impairment of the ability to form high molecular weight AHF from the building stones—the protein chains. Since it is only the high molecular weight AHF that is precipitated with rabbit antiserum, the antigen content will be small or absent. Secondarily, a feedback inhibition of the synthesis of low molecular weight AHF occurs (Figure 2).

VWD with a normal content of AHF-related antigen has also been described (25). One might imagine several explanations:

(*a*) There is X-chromosomal heredity with a defect in the sequence of the primary structure which is significant for VWF activity. Also the AHF activity is affected. The polymerization is normal (as in hemophilia A) because the autosomal genes are intact. This variant might also be regarded as a variant of hemophilia A with a prolonged bleeding time and decreased platelet adhesiveness (25), but clinically it resembles VWD rather than hemophilia A.

(*b*) There is autosomal heredity in which polymerization may occur, but in an abnormal way, with the effect that the quaternary structure is different, resulting in reduction of both the VWF and the AHF activity. The structure is, however, so similar to normal that the molecule reacts with rabbit antiserum against AHF-related protein.

How and where does the VWF act in vivo? As mentioned, the factor occurs in the plasma, but with the aid of immunofluorescence it can be demonstrated also in

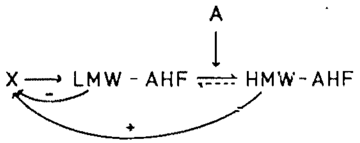


Figure 2 Hypothetical pathway for the synthesis of AHF and VWF. *X:* X-chromosomal gene, *A:* autosomal gene, *LMW-AHF:* low molecular weight AHF, *HMW-AHF:* high molecular weight AHF.

the intima of the blood vessels (39). The protein may actually be synthesized in endothelial cells (40). Patients with severe VWD lack AHF-VWF protein not only in the plasma, but also in the vessel walls (41). In such patients infusion of cryoprecipitate is followed by an increase in the AHF-VWF antigen in the plasma, while the antigen is still not demonstrable in the endothelium (Holmberg, unpublished). This finding suggests that the occurrence of AHF-VWF protein in the intima is due to a true synthesis in the vessel wall and not simply to adsorption of the protein to the endothelial surface. It is not known whether the plasma protein or the vessel protein is of greater importance in primary hemostasis.

The interaction between AHF-VWF protein and platelets in primary hemostasis is still unknown. Some interesting observations have, however, been made. Purified bovine and porcine high molecular weight AHF-related proteins aggregate human platelets in vitro (42). Human AHF-VWF protein can be modified in such a way as to aggregate platelets. The antibiotic ristocetin aggregates platelets in vitro in normal persons, but generally not in patients with VWD (43). Ristocetin has no effect on isolated platelets, but requires the presence of the high molecular weight AHF-VWF protein to exert its effect (44).

The effect of the enzyme neuraminidase on platelet aggregation also seems to be related to the AHF-VWF protein. Neuraminidase aggregates rabbit platelets in the presence of a plasma factor (45), and increases aggregation of human platelets with inter alia ADP (46). Vermylen et al (47) showed that human AHF-VWF protein, preincubated with neuraminidase, spontaneously aggregates human platelets in platelet-rich plasma.

GENETICS

VWD seems to be one of the most common hemorrhagic diatheses. Silwer (48) estimated the frequency in Sweden at 1:32,000 inhabitants. However, mild cases may remain undiagnosed.

Most authors agree that the disease is inherited by an autosomal and dominant gene (10, 49). Of 40 families with VWD, Silwer (48) found the expected pattern of autosomal dominant inheritance in 27 (Figure 3).

In single published cases, however, the transmission might have been different (25, 48, 50). Some authors have shown an increased frequency of consanguineous marriages, others a high frequency of sporadic cases (51). The disease seems to be more common among women than among men. This is probably because women with VWD so often have menorrhagia, which leads to the detection of the disease.

Assuming that the inheritance is autosomal dominant, it is obvious that its penetrance and expressivity vary considerably (48, 51, 52). Yet in Silwer's material substantial differences in expressivity were uncommon. He found the penetrance to be 73–90%, and estimated the risk of a child of an affected person inheriting the disease to be 40% or less. The risk of the disease being severe was, however, regarded as 5% or less.

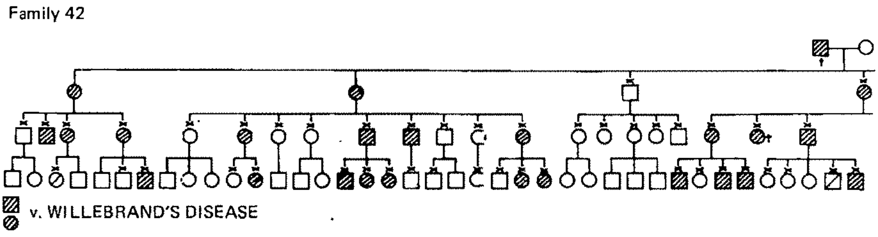


Figure 3 Pedigree illustrating autosomal and dominant inheritance of v. Willebrand's disease.

SYMPTOMS

The symptoms in VWD vary widely. In severe VWD the bleedings may be life-threatening, while in mild VWD they may be so mild that the disease remains concealed.

Nose bleeding is the most common symptom of VWD and occurs in 60–80% of affected patients (Silwer, 48). Bleeding from the oral cavity is also fairly common; in children, traumatic bleeding of the lips, tongue, and oral cavity are characteristic. In contrast with what is seen in hemophilia, deep subcutaneous and intramuscular hematomas are uncommon in VWD. Very typical, however, are prolonged bleedings from trivial wounds.

The most prominent bleeding symptom in women with VWD is probably menorrhagia. Menstrual disorders occur in some 50–75% of such women. The symptom is most severe during the first few years after menarche, and has occasionally been life-threatening and required hysterectomy or roentgen castration. Menorrhagia is also often a difficult problem in VWD characterized as mild by laboratory tests. Another type of bleeding from the female genital tract, which has not received much attention, is corpus luteum bleeding. Typical cases are characterized by intermenstrual, more or less severe, abdominal pain and often by a low Hb because of intra-abdominal bleeding.

Delivery in women with VWD is often, though not always, associated with heavy blood loss. In some women with VWD the AHF level rises during pregnancy, which lessens the risk of bleeding. The AHF, however, does not always rise, and especially not in women with an initially very low level. Even if no bleeding occurs at parturition, the patient must be observed for at least a week post partum because the AHF level falls after delivery.

In most published series of VWD the frequency of gastrointestinal bleeding is relatively low, usually 10–15%. It is, however, striking that when such bleeding occurs it is usually very severe. In most cases the patients appear to bleed diffusely from the mucosa. Roentgen examination rarely reveals any local source of bleeding, such as gastric or duodenal ulcer.

The frequency of hematuria has been given as 2–7.5%. The bleeding is rarely so severe as to require specific treatment.

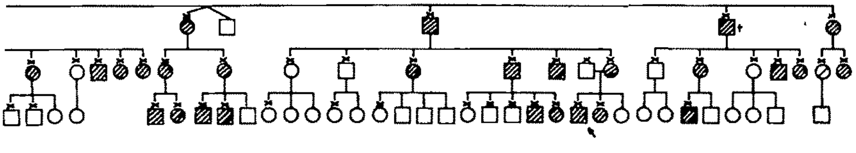


Figure 3 (Continued)

Joint bleedings are not very common in VWD, in contrast to hemophilia, but do occur in patients with severe disease. In a Swedish series of 32 cases of severe VWD, half of the patients had hemarthroses and seven chronic arthropathy.

Postoperative bleeding is a common and serious symptom in VWD. Tooth extractions are often followed by prolonged oozing, unless adequate precautions have been taken. Operations in the region of the ear, nose, or throat such as tonsillectomy and adenoidectomy are often associated with prolonged serious bleeding. This is true also for abdominal operations, although it has sometimes been possible to perform an appendectomy without special precautions in patients with mild VWD. In all operations the bleedings may increase or recur several days or even weeks after the operation.

DIAGNOSIS

The manifestations of bleeding, like the laboratory findings in VWD, tend to vary from time to time (1, 53). The symptoms often decrease in severity with increasing age. The diagnosis of VWD must therefore be made jointly on the basis of the patient's history and several different laboratory examinations. We have used the following criteria for the diagnosis of VWD: 1. bleeding in the patient's history, 2. positive heredity, 3. prolonged bleeding time, 4. decreased AHF activity (60%), 5. decreased platelet adhesiveness, and 6. decreased AHF-related antigen.

We also think it possible to diagnose VWD in those cases where (*a*) relatives have not been examined but the patient meets the requirements given under points 1 and 3–6, and (*b*) typical disease in the family, but not all findings under points 3–6, are positive.

The above-mentioned variations in the laboratory values must always be borne in mind. In patients with VWD, as in normals, the AHF activity fluctuates widely (51, 53). This applies particularly to mild and moderate cases of VWD. The AHF activity also increases with increasing age (48).

The bleeding time also tends to become shorter with increasing age. Usually Duke's or Ivy's method is used. Duke's method is rather insensitive and has often been reported to fluctuate. In severe VWD, Duke's bleeding time is, however, nearly always markedly prolonged. The Ivy bleeding time, and especially the modification by Borchgrevink & Waaler (54), is a much more sensitive method. It is prolonged also in mild forms of the disease, where the Duke bleeding time is normal (55, 56).

Measurement of platelet adhesiveness is now widely used to diagnose VWD. Salzman's method is used the most, and its value is well documented (19, 48, 51–53). Patients with VWD have, as a rule, low values, below 20%, but there is a certain overlap between normal persons and those with VWD. We therefore feel that reduced platelet adhesiveness alone cannot be taken as evidence of a symptom-free person having VWD or being a carrier of the disease, even if typical disease is known in the family.

A firm diagnosis of VWD cannot be made before one has studied the effect of an infusion of AHF concentrate (fraction I–0 or cryoprecipitate), at least in the proband in each family. The Duke bleeding time should become shorter, while the Ivy bleeding time is less responsive. The progressive increase in AHF activity is diagnostic, but may sometimes be difficult to demonstrate in patients with mild VWD (51, 57, 58).

Immunological determination of AHF-related antigen is often useful, especially in family investigations. The determination requires a specific antiserum. Most patients with VWD have a low level of AHF-related antigen (23, 25).

VARIANTS OF VWD

Several authors have described VWD or VWD-like clinical pictures in which the content of AHF-related antigen has been normal. In such cases it must be assumed that the protein is defective both in respect to its AHF and VWF functions. Holmberg & Nilsson (25) described several VWD families in which all the affected members had a normal content of AHF-related antigen. In at least some of the families the inheritance was X-chromosomal. Similar patients with a normal content of antigen have since been described by several authors (59–62).

Another variant of VWD-like disease has been observed by Nilsson & Cronberg (63). It is characterized by a prolonged bleeding time, decreased platelet adhesiveness, normal AHF activity, but decreased amount of AHF antigen (Holmberg, unpublished).

Decreased ristocetin aggregation of platelets in platelet-rich plasma is typical of VWD, but not obligatory (64). There are probably many different patterns of the relation among AHF activity, AHF-related antigen, and ristocetin aggregation (62).

From these variants one must separate those with simultaneous occurrence of VWD and other bleeding disorders, e.g. platelet defects (65–67) and classical hemophilia (68, 69). A so-called acquired VW syndrome has been seen, especially in systemic lupus erythematosus and myeloma (57, 70, 71).

TREATMENT

The discovery that VWD is due to lack of a plasma factor opened up a new approach to treatment: administration of fresh plasma. Good therapeutic and prophylactic results have been achieved, but in severe VWD it has often proved to be less suitable because administration of sufficient volumes to secure hemostasis can cause a strain on the circulation.

The first acceptable concentrate was the Swedish fraction I–0 (10). This preparation has been used with success for almost 20 yr, and has proved capable of controlling all types of bleeding in VWD. Cryoprecipitate has also been reported to have a good effect.

Experience has shown that the bleedings are controlled when the Duke bleeding time is normalized. The effect on the Duke bleeding time is shorter than that on the AHF activity level (see above), thus bleeding may occur even if the AHF level is high.

The specific substitution therapy is the keystone of treatment. The most serious and common complication in the treatment with plasma concentrate is serum hepatitis. It is, however, fairly rare and only mild with fraction I–0.

Oral contraceptives have been found to have a considerable depressive effect on menorrhagia. This effect can probably be ascribed mainly to a local action of the hormone on the endometrium.

Therapy and prophylaxis with inhibitors of fibrinolysis, epsilon aminocaproic acid (EACA) and tranexamic acid (AMCA), have proved valuable in menorrhagia in patients who cannot use oral contraceptives. Inhibitors of fibrinolysis should also be used with specific treatment for bleeding from the ear, nose, and throat and in patients about to undergo tooth extraction (72). We recommend a large dose of AHF concentrate and AMCA intravenously before extraction of a tooth. The patient should afterwards be given AMCA orally for one week. As a rule, no further substitution therapy is necessary.

Certain drugs are contraindicated in VWD. Acetylsalicylic acid prolongs the bleeding time and accentuates the tendency to bleed. This also holds for dextran, which reduces platelet adhesion. It should also be observed that intramuscular injections are likely to cause large hematomas.

SUMMARY

The common inheritable hemorrhagic disorder, von Willebrand's disease, has been known for only about 50 yr. During these years it has been well established that the cause of the disease is a deficiency of a plasma protein with remarkable biological properties. Future research should focus on details of the molecular structure of this protein. It is hoped that basic science will help to elucidate those aspects of von Willebrand's disease that are not yet properly understood.

Literature Cited

1. von Willebrand, E. A. 1926. *Finska Läkaresällskapets Handlingar* 67(2): 7–112
2. Minot, G. R. 1928. A familial hemorrhagic condition associated with prolongation of the bleeding time. *Am. J. Med. Sci.* 175:301–6
3. Jürgens, R. 1937. Die erblichen Thrombopathien. *Ergeb. Inn. Med. Kinderheilk.* 53:795–826
4. Mac Farlane, R. G. 1941. Critical review: the mechanism of haemostasis. *Quart. J. Med.* 10:1–30
5. Estren, S., Sanchez Médal, L., Dameshek, W. 1946. Pseudohemophilia. *Blood* 1:504–33
6. Alexander, B., Goldstein, R. 1953. Dual hemostatic defect in pseudohemophilia. *J. Clin. Invest.* 32:551 (Abstr.)
7. Schulman, I., Smith, C. H., Erlandson, M., Fort, E., Lee, R. E. 1956. Vascular hemophilia. A familial hemorrhagic disease in males and females characterized by combined antihemophilic globulin deficiency and vascular abnormality. *Pediatrics* 18:347–61
8. Singer, K., Ramot, B. 1956. Pseudohemophilia type B. *Arch. Intern. Med.* 97:715–25
9. Nilsson, I. M., Blombäck, M., Jorpes, E., Blombäck, B., Johansson, S.-A. 1957. v. Willebrand's disease and its correction with human plasma fraction I–0. *Acta Med. Scand.* 159:179–88
10. Nilsson, I. M., Blombäck, M., v. Francken, I. 1957. On an inherited autosomal hemorrhagic diathesis with antihemophilic globulin (AHG) deficiency and prolonged bleeding time. *Acta Med. Scand.* 159:35–57
11. Blombäck, B., Blombäck, M. 1956. Purification of human and bovine fibrinogen. *Ark. Kemi* 10:415–33
12. Nilsson, I. M., Blombäck, M., Blombäck, B. 1959. v. Willebrand's disease in Sweden. Its pathogenesis and treatment. *Acta Med. Scand.* 164:263–78
13. Cornu, P., Larrieu, M.-J., Caen, J., Bernard, J. 1960. Données nouvelles concernant la maladie de Willebrand (angiohémophilie). *Rev. Fr. Etud. Clin. Biol.* 5:614–20
14. van Creveld, S., Mochtar, I. A. 1960. von Willebrand's disease—a plasma deficiency cause of the prolonged bleeding time. *Ann. Pediat.* 194:37–46
15. Bennett, E., Dormandy, K. 1966. Pool's cryoprecipitate and exhausted plasma in the treatment of von Willebrand's disease and factor XI-deficiency. *Lancet* 2:731–32
16. Borchgrevink, Chr. F. 1961. Platelet adhesion in vivo in patients with bleeding disorders. *Acta Med. Scand.* 170:231–43
17. Hellem, A. J. 1960. The adhesiveness of human blood platelets in vitro. *Scand. J. Clin. Lab. Invest. Suppl.* 51:1–110
18. Cronberg, S., Nilsson, I. M., Silwer, J. 1966. Studies on the platelet adhesiveness in von Willebrand's disease. *Acta Med. Scand.* 180:43–54
19. Salzman, E. W. 1963. Measurement of platelet adhesiveness. A simple in vitro technique demonstrating an abnormality in von Willebrand's disease. *J. Lab. Clin. Med.* 62:724–35
20. Bowie, E. J. W., Owen, C. A., Thompson, J. H., Didisheim, P. 1969. Platelet adhesiveness in von Willebrand's disease. *Am. J. Clin. Pathol.* 52:69–77
21. Hellem, A. J. 1970. Platelet adhesiveness in von Willebrand's disease. *Scand. J. Haematol.* 7:374–82
22. Meyer, D., Larrieu, M.-J. 1970. von Willebrand factor and platelet adhesiveness. *J. Clin. Pathol.* 23:228–31
23. Zimmerman, T. S., Ratnoff, O. D., Powell, A. E. 1971. Immunologic differentiation of classic hemophilia (factor VIII deficiency) and von Willebrand's disease. *J. Clin. Invest.* 50: 244–54
24. Holmberg, L., Nilsson, I. M. 1973. Immunologic studies in haemophilia A. *Scand. J. Haematol.* 10:12–16
25. Holmberg, L., Nilsson, I. M. 1972. Genetic variants of von Willebrand's disease. *Brit. Med. J.* 3:317–20
26. van Mourik, J. A., Mochtar, I. A. 1970. Purification of human antihemophilic factor (factor VIII) by gel chromatography. *Biochim. Biophys. Acta* 221: 677–79
27. Bouma, B. N., Wiegerinck, Y., Sixma, J. J., van Mourik, J. A., Mochtar, I. A. 1972. Immunological characterization of purified anti-haemophilic factor A (factor VIII) which corrects abnormal platelet retention in von Willebrand's disease. *Nature New Biol.* 236:104–6
28. Shapiro, G. A., Andersen, J. C., Pizzo, S. V., McKee, P. A. 1973. The subunit structure of normal and hemophilic factor VIII. *J. Clin. Invest.* 52:2198–2210
29. Holmberg, L., Nilsson, I. M. 1974. AHF related protein in clinical praxis. *Scand. J. Haematol.* 12:221–31

30. Rick, M. E., Hoyer, L. W. 1973. Immunologic studies of antihemophilic factor (AHF, factor VIII). V. Immunologic properties of AHF subunits produced by salt dissociation. *Blood* 42:737–47
31. Owen, W. G., Wagner, R. H. 1972. Antihemophilic factor: separation of an active fragment following dissociation by salts or detergents. *Thromb. Diath. Haemorrh.* 27:502–15
32. Weiss, H. J., Hoyer, L. W. 1973. von Willebrand factor: Dissociation from antihemophilic factor procoagulant activity. *Science* 182:1149–51
33. Zimmerman, T. S., Edgington, T. S. 1973. Factor VIII coagulant activity and factor VIII-like antigen: independent molecular entities. *J. Exp. Med.* 138:1015–20
34. van Mourik, J. A., Bouma, B. N., LaBruyère, W. T., de Graaf, S., Mochtar, I. A. 1974. Factor VIII, a series of homologous oligomers and a complex of two proteins. *Thromb. Res.* 4:155–64
35. Bennett, B., Ratnoff, O. D., Levin, J. 1972. Immunologic studies in von Willebrand's disease. *J. Clin. Invest.* 51: 2597–2601
36. Holmberg, L., Nilsson, I. M. 1973. Two genetic variants of von Willebrand's disease. *N. Engl. J. Med.* 288:595–98
37. Bloom, A. L., Peake, I. R., Giddings, J. C. 1973. The presence and reactions of high and lower-molecular-weight procoagulant and Factor VIII in the plasma of patients with von Willebrand's disease after treatment: significance for a structural hypothesis for Factor VIII. *Thromb. Res.* 3:389–404
38. Bennett, B., Forman, W. B., Ratnoff, O. D. 1973. Studies on the nature of antihemophilic factor (factor VIII). *J. Clin. Invest.* 52:2191–97
39. Bloom, A. L., Giddings, J. C., Wilks, C. J. 1973. Factor VIII on the vascular intima: possible importance in haemostasis and thrombosis. *Nature New Biol.* 241:217–19
40. Jaffe, E. A., Hoyer, L. W., Nachman, R. L. 1973. Synthesis of antihemophilic factor antigen by cultured human endothelial cells. *J. Clin. Invest.* 52:2757–64
41. Holmberg, L., Mannucci, P. M., Turesson, I., Ruggeri, Z. M., Nilsson, I. M. 1974. Factor VIII antigen in the vessel wall in von Willebrand's disease and haemophilia A. *Scand. J. Haematol.* 13: 33–38
42. Forbes, C. D., Prentice, C. R. M. 1973. Aggregation of human platelets by purified porcine and bovine antihaemophilic factor. *Nature New Biol.* 241:149–50
43. Howard, M. A., Firkin, B. G. 1971. Ristocetin—A new tool in the investigation of platelet aggregation. *Thromb. Diath. Haemorrh.* 26:362–69
44. Weiss, H. J., Rogers, J., Brand, H. 1973. Defective ristocetin-induced platelet aggregation in von Willebrand's disease and its correction by factor VIII. *J. Clin. Invest.* 52:2697–2707
45. Hovig, T. 1965. The effect of various enzymes on the ultrastructure, aggregation and clot retraction ability of rabbit blood platelets. *Thromb. Diath. Haemorrh.* 13:84–113
46. Davis, J. W., Yue, K. T. N., Phillips, P. E. 1972. The effect of neuraminidase on platelet aggregation induced by ADP, norepinephrine, collagen or serotonin. *Thromb. Diath. Haemorrh.* 28:221–27
47. Vermylen, J., Donati, M. B., de Gaetano, G., Verstraete, M. 1973. Aggregation of human platelets by bovine or human factor VIII: role of carbohydrate side chains. *Nature* 244:167–68
48. Silwer, J. 1973. von Willebrand's disease in Sweden. *Acta Paediat. Scand. Suppl.* 238:1–159
49. Barrow, E. M., Graham, J. B. 1964. von Willebrand's disease. *Progr. Hematol.* 4:203–21
50. Veltkamp, J. J., Tilburg, N. H. 1973. Detection of heterozygotes for recessive von Willebrand's disease by the assay of antihemophilic-factor-like antigen. *N. Engl. J. Med.* 289:882–85
51. Strauss, H. S., Bloom, G. E. 1965. von Willebrand's disease. Use of a platelet-adhesiveness test in diagnosis and family investigation. *N. Engl. J. Med.* 273:171–81
52. Larrieu, M. J. et al 1968. Congenital bleeding disorders with long bleeding time and normal platelet count. II. von Willebrand's disease (report of thirty-seven patients). *Am. J. Med.* 45:354–72
53. Abildgaard, C. F. et al 1968. von Willebrand's disease: A comparative study of diagnostic tests. *Pediatrics* 73:355–63
54. Borchgrevink, Ch. F., Waaler, B. A. 1958. The secondary bleeding time. A new method for the differentiation of hemorrhagic diseases. *Acta Med. Scand.* 162:362–74
55. Nilsson, I. M., Magnusson, S., Borchgrevink, C. 1963. The Duke and Ivy methods for determination of the bleeding time. *Thromb. Diath. Haemorrh.* 10:223–34

56. Silwer, J., Nilsson, I. M. 1964. On a Swedish family with 51 members affected by von Willebrand's disease. *Acta Med. Scand.* 175:627–43
57. Ingram, G. I. C., Kingston, P. J., Leslie, J., Bowie, E. J. W. 1971. Four cases of acquired von Willebrand's syndrome. *Brit. J. Haematol.* 21:189–99
58. Veltkamp, J. J., van Tilburg, N. H. 1974. "Autosomal haemophilia": a variant of von Willebrand's disease. *Brit. J. Haematol.* 26:141–52
59. Bouma, B. et al 1973. Factor-VIII antigen and platelet retention in a glass bead column. *Brit. J. Haematol.* 25:645–56
60. Kernoff, P. B. A., Gruson, R., Rizza, C. R. 1974. A variant of factor VIII related antigen. *Brit. J. Haematol.* 26:435–40
61. Thomson, C., Forbes, C. D., Prentice, C. R. M. 1974. Abnormal factor VIII in von Willebrand's disease. *Scot. Med. J.* 19:58
62. Firkin, B., Koutts, J. 1974. VIII-related antigen and von Willebrand factor. *Lancet* 1:803–4
63. Nilsson, I. M., Cronberg, S. 1968. A severe haemorrhagic disorder with prolonged bleeding time due to a plasma defect but with normal factor VIII. *Acta Med. Scand.* 184:181–86
64. Howard, M. A., Sawers, R. J., Firkin, B. G. 1973. Ristocetin: a means of differentiating von Willebrand's disease into two groups. *Blood* 41:687–90
65. Blackburn, E. K., Macfie, J. M., Monaghan, J. H., Page, A. P. M. 1961. Antihaemophilic factor deficiency, capillary defect of von Willebrand type, and idiopathic thrombocytopenia occurring in one family. *J. Clin. Pathol.* 14:540–42
66. Quick, A. J. 1965. Hereditary thrombopathic thrombocytopenia and Minot-v Willebrand syndrome: probable coexistence in a family. *Am. J. Med. Sci.* 250:1–7
67. Corder, M. P., Culp, N. W., O'Neil, B. Jr. 1973. Familial occurrence of von Willebrand's disease, thrombocytopenia, and severe gastrointestinal bleeding. *Am. J. Med. Sci.* 265:219–23
68. Quick, A. J., Adlam, R. T. 1963. Coexistence of von Willebrand's disease and hemophilia in a family. *J. Am. Med. Assoc.* 185:635–39
69. Geiger, M. T., Rath, Ch. E. 1963. Occurrence of two hemorrhagic disorders with antihemophilic factor (AHF) deficiency in the same family: classical hemophilia and von Willebrand's disease. *J. Lab. Clin. Med.* 61:424–36
70. Simone, J. V., Cornet, J. A., Abildgaard, C. F. 1968. Acquired von Willebrand's syndrome in systemic lupus erythematosus. *Blood* 31:806–12
71. Mant, M. J. et al 1973. von Willebrand's syndrome presenting as an acquired bleeding disorder in association with a monoclonal gammopathy. *Blood* 42:429–36
72. Björlin, G., Nilsson, I. M. 1973. Tooth extractions in hemophiliacs after administration of a single dose of factor VIII or factor IX concentrate supplemented with AMCA. *Oral Surg. Oral Med. Oral Path.* 36:482–89

Copyright 1975. All rights reserved

THE MINIMAL BRAIN DYSFUNCTION SYNDROME

❖7130

Paul H. Wender, M.D.
Department of Psychiatry, University of Utah College of Medicine, Salt Lake City, Utah 84132

The phrase minimal brain dysfunction (MBD) is one of a bewildering number of terms applied to a behavioral and cognitive syndrome of childhood. A limited list of these partial synonyms would include: postencephalitic behavior disorder, hyperkinesis, hyperactivity, the hyperactive child syndrome, minimal brain damage, minimal cerebral dysfunction, specific learning disabilities, etc. MBD is probably the least objectionable term in that it implies no specific etiology, no pathognomonic signs and symptoms, and no homogeneity of cause, course, or treatment. A few grammarians have objected to the use of the word "minimal" since some individuals with the disorder may have serious handicaps. This criticism is valid, but barring the discovery of an 18th century physician whose eponym may henceforth be used to denote the syndrome, MBD is the term that is the best of a number of bad alternatives.

MBD is a syndrome of considerable practical importance. Hard, reliable diagnostic categorization is not available in child psychiatry. However, from the impressions of clinicians working in varying settings, it would appear that the syndrome constitutes the single most common cause of chronic behavioral problems in the pre-adolescent age group. Important, too, is the fact that of all the chronic psychological and psychiatric problems of childhood, MBD is clearly the syndrome which most obviously responds to relatively minimal intervention. These facts necessitate a broader recognition of the signs of the syndrome and the modalities most useful in its treatment.

CHARACTERISTICS OF MBD

Before specifying the agreed-upon characteristics of MBD children, it is important to emphasize the epistemological difficulties in delimiting the boundaries of the syndrome. As in the case of all psychiatric diagnostic categories, we are at present unable to turn to a pathologist, biochemist, or geneticist to specify a group of individuals whose signs and symptoms we may then study. (Compare rheumatic fever in which we were able to link the various clinical manifestations together in a common syndrome only when the pathologist was able to specify underlying and defining tissue and immunological characteristics for the syndrome.) Accordingly, we can only specify signs and symptoms which cluster together. We can not specify whether certain signs by themselves are representative of the diagnostic category. A further factor in establishing the boundaries of the syndrome is that clinical reports from different settings have utilized different (and not defined) clinical populations. The source of referral influences the population evaluated, which in turn influences the prevalences, prognoses, and responses to treatment reported. For this reason, no attempt will be made to provide exact definitions. Similarly, it would be misleading to give exact percentages; only order of magnitude figures will be provided. In other words, the boundaries are blurred.

With this caveat in mind, let us turn to the typical abnormalities of MBD children. It is important to emphasize that the behavior of children with this disorder is, in general, not qualitatively different from that of the normal child. The differences between the affected child and the normal child are in the intensity, persistence, and clustering of his signs and symptoms. Further complicating the picture is the fact that some of the salient characteristics tend to change as the child matures. Since there is considerable disagreement as to the boundaries of the syndrome, I shall here consider only those attributes that are found in classical instances.

There are two major areas of dysfunction found in MBD children: (*a*) behavioral problems, and (*b*) perceptual-cognitive problems. These two areas of dysfunction occur in three possible combinations: (*a*) children with behavioral difficulties alone, (*b*) children with both perceptual and cognitive difficulties, (*c*) children with difficulties in both areas. The discrimination into the three subgroups is of practical importance in that the prognosis and therapeutic modalities are determined by which areas of dysfunction the child manifests.

Behavioral Abnormalities (1, 2)

MOTOR BEHAVIOR The two most characteristic alterations of motor behavior in MBD children are an increased level of activity (hence the terms "hyperactivity" and "hyperkinesis") and impaired coordination. With regard to the former, the mother is apt to report that the child had been excessively active from infancy, that as he grew older he was always on the go, ran rather than walked, and was always into things. In kindergarten and elementary school, his teachers usually observed that the child was unable to sit still at his desk, frequently getting up from his seat, walking around the room, humming, whistling, tapping, and generally bothering his classmates and teacher. During adolescence this overactivity, when present, may disappear; it is this frequent maturational alteration that has led to the erroneous observation that all of the signs and symptoms of MBD diminish at that time. Attempts to document this increased activity by objective techniques have been successful with some methods but not with others. Increased wiggling in a classroom seat does take place (3) but increased ambulation, as measured by a pedometer (4) does not. It is probably the case that MBD children are not universally excessively active but that their activity is inappropriate. In structured settings, their behavior is frequently characterized by rapid changes of attention from one activity to another, which may appear to an observer as restlessness, but is not necessarily equivalent to increased general activity. Excessive motor activity, when present, often diminishes or disappears around puberty. It is important to emphasize that hyperactivity is not a sine qua non of minimal brain dysfunction. Some children with all the other signs and symptoms of MBD are normally active or hypoactive.

Another area of motor abnormality, present in approximately one half of MBD children, is some degree of motor incoordination. The population distribution of MBD children may be bimodal since some MBD children are reported by their parents to have been usually advanced in their rate of motor development and to currently manifest excellent motor coordination. When such incoordination is present, it may manifest itself in fine motor performance, hand-eye coordination, or balance. Such function is often assessed as well by history as by a neurological examination. Children in the first category will be reported to have had difficulties in learning to button buttons, zip zippers, tie shoelaces, color, cut with scissors, and eventually, write. Children with hand-eye incoordination manifest, of course, poor ability in catching, hitting, and sometimes throwing a ball; usually the diagnosis is made by the child's peer group who are expert neurological diagnosticians in this respect and are apt to chose the MBD child last in games. Finally, gross balance may be impaired and the child may have had difficulty learning to roller skate or ride a two-wheeled bicycle.

ATTENTIONAL DIFFICULTIES Whether "attention" should be placed under perceptual and cognitive characteristics or under general characteristics is moot. In view of the fact that many MBD children show attentional deficits with no other perceptual or cognitive characteristics, it is probably a good idea to consider attentional deficits separately. Attentional difficulties are present in most, if not all, MBD children. This characteristic is first noted by the mother in terms of the child's inability or unwillingness to persist for very long in any activities; in the most extreme instances he may be described as having been unable to engage in any activity for an extended period of time. At school his teacher will report that he has a short attention span (in the current educational jargon), is distractible, and cannot stick with academic tasks for a sufficient period of time. Attention, like the other variables we are considering, occurs along a continuum. Many children fail to display the eager interest in school subjects which their teachers would wish. The question is again one of degree. In clear-cut instances of MBD, teachers and tutors will report that the child is able to stay with his work only when he receives constant one-to-one attention and reinforcement. Of considerable importance is the fact that the attentional deficits, unlike increased motor activity, do not seem to disappear at puberty, although they may assume a more muted form.

IMPULSE CONTROL Another striking, and sometimes the most distressing, characteristic of many MBD children is an impairment of impulse control. Together with academic underachievement (to be discussed later), lack of impulse control probably constitutes the most common referring complaint. The complaints differ as a function of the observer and the setting. Both parents and teachers agree that the child is and has been difficult to control from without and seems to manifest little self-control. The manifestations of impulsivity vary as a function of age: as a toddler the mother reports that the child was difficult to toilet train, that he was difficult to keep from running out into the street, and that he was likely to experience more than his share of accidents (including "accidental" poisonings). As a school-age child he had difficulty in refraining from the frequent temptations of that age period, taking things that were not his, playing with matches, etc. If the difficulties persist into adolescence, they take the form of impulsiveness of that period: stealing, delinquency, sexual "acting out." It is probable that children with MBD are overrepresented among populations of juvenile delinquents, as suggested by both retrospective and anterospective studies.

INTERPERSONAL RELATIONS The child with MBD appears to have certain characteristic forms of interpersonal behavior with both his peers and adults: He tends to be extroverted, independent, and controlling. With adults he tends to be resistant to social demands and expectations, and a frequent parental complaint is that they have to treat him more firmly and authoritatively than their other children if he is to conform, even minimally to familial rules. In extreme instances he is described as being stubborn, negativistic, and refractory to all ordinary disciplinary measures. Typical parental comments include: "You have to tell him things a thousand times . . . you can't get through to him . . . punishment rolls right off his back . . . you have to stay with him at all times." The refractoriness to discipline of which parents frequently complain is often attributed to inconsistency and other parental failings. While there is no doubt that inadequate firmness and limit setting may aggravate the problems of MBD children, exactly the same problems are seen in families where the parents have managed their other children quite adequately. For this reason, it is of practical usefulness to consider the diagnosis of MBD in a persistently disobedient child from a fairly well ordered family. In the past, it has been unfortunately common to attribute such difficulties in a child to inadequate parental treatment. Not only is this often unfair but it is often inaccurate and often obscures the diagnosis and consequently the treatment of the affected child. With peers as with adults, the MBD child's interpersonal style is likely to lose friends and alienate people. Children are apt to find him bossy, aggressive, and unwilling to follow the rules of the group. He is therefore frequently rejected and consequently described as friendless. It is of diagnostic importance to distinguish between the MBD child who has no friends because of his offensiveness and the schizoid child who has usually been socially avoidant and shy from an early age.

"EMOTIONAL" ABNORMALITIES The ill-defined term "emotional" will be used to refer to behavioral alterations which include: altered sensory thresholds, increased lability, altered reactivity, increased aggressiveness, and dysphorias (a nonspecific term for bad mood, including depression).

First, many MBD children's pain threshholds tend to be high. This may be fortunate in that his impulsivity seems to make him liable to greater than his share of the minor traumata of childhood. Mood fluctuations from day to day and within the day are often, but not invariably, reported by parents. Short-term lability is very common. Many children have a "short fuse" or "low boiling point," reacting easily and persistently to everyday frustrations. Temper tantrums are frequent and continue later than would be expected. Children with MBD are frequently described as "angry;" most are not. They do have a lower threshhold for loss of control, as noted above, but do not tend to be cruel. Excessive cruelty—to animals or other children—is not typical nor is a preoccupation with aggression or violence. Sadism and macabre preoccupations are more characteristic of the borderline schizophrenic child. The principal dysphoric characteristics are anhedonia (diminished ability to experience pleasure), depression, and low self-esteem. In some instances, anhedonia has been present from infancy and the parents will report that the child has never really been satisfied with anything or obtained long-lasting pleasure from toys or accomplishments. Such attributes are often erroneously interpreted as the result of spoiling, but a long history suggests another etiology. By the time MBD children reach the age of relatively accurate self reports, many will state that they feel depressed or hold a low opinion of themselves; many will appear to be this way but will deny such feelings. In the latter instance, the child's low self-esteem may be superficially concealed by a boastful attitude but observant parents and teachers will often report comments that reveal that the child perceives himself as inadequate, stupid, or bad.

Perceptual-Cognitive Difficulties

As previously stated, this class of problems can occur either with or independently of the behavioral problems already described. Historically, some children of normal intelligence were found to have extreme difficulty in learning basic scholastic skills, particularly reading, spelling, and, in some instances, arithmetic calculations. This group of children, identified by their impaired cognitive performance, was labeled by a variety of phrases including dyslexia, specific learning disability, etc. Educational observers focused on the cognitive rather than the behavioral difficulties, tended to overlook them, and hence the overlap between behavioral and perceptual-cognitive problems was ignored. Only recently has the overlap in problem areas been noted.

In an attempt to understand the underlying psychological dysfunctions in this group of children, a number of empirical tests were administered. This group of test abnormalities, which MBD children with perceptual-cognitive problems show, are only correlates and not explanations: they may be only other manifestations of the true causes of the observed defects rather than the cause of these defects. Among the perceptual-cognitive difficulties frequently found both in vivo and on tests are: problems of orientation in space, as manifested by rotation of copied geometrical designs on the Bender Gestalt Test, right-left disorientation, transpositions of letters (reversal of positions of letters), and reversals of letters themselves in reading and writing; difficulties in auditory discrimination (so that similar phonemes are confused); difficulties in auditory synthesis (so that the successively, and correctly, sounded out phonemes can not be grouped together to form a word); difficulties in auditory and visual memory; and difficulty in transferring information from one sensory modality (e.g. auditory) to another (e.g. visual). Such psychological impairments are not assessed on routine psychological tests but are measured on tests employed by special educators. As mentioned, these perceptual-cognitive difficulties may occur in children at all levels of intellectual function. Since at least some of these functions appear to be related to what is generally deemed intelligence, it may seem paradoxical that such difficulties may coexist with normal intelligence as typically measured. This contradiction is a consequence of the construction of intelligence tests; many older tests did employ measures of spatial relations, and MBD children with these psychological characteristics would test as impaired in such functions. Indeed, these cognitive problems can be perceived as focal areas of intellectual impairment and it is only, as noted, an artifact of the commonly used tests that these children are not penalized for these deficits.

Attempts have been made to divide these "learning disabled" MBD children into homogeneous categories. One such attempt, which jibes with clinical impression, is described by Boder (5). Operationally, one finds two distinct subgroups (and a much larger group of children manifesting both sets of characteristics). First, there are children who are able to read only words that they have previously learned. These words will be spelled perfectly. When given new words, they are unable to sound them out. When asked to spell words that they have not previously encountered, their spelling will be bizarre. They have not mastered the rules of phonetic spelling. Presumably, these children have a good visual memory which enables them to recognize and reproduce words previously encountered. They are unable to sound out new words or to transfer heard (new) sounds to symbols. They are often said to have auditory problems. The second group, those with visual problems, can sound out words they have never previously seen; they have mastered the phonetic technique. When asked to spell words that they have either previously learned or that they have never spelled before, they reproduce these words phonetically. They reproduce nonphonetically spelled words phonetically. They appear to have a deficient visual memory and go awry with words that are spelled differently from the way they sound. As is usually the cases, the pure types are less common than the mixed—children who have both auditory and visual problems. Both groups of children are reported to have the other (listed) perceptual-cognitive problems described before.

LEARNING DIFFICULTIES A large fraction of MBD children manifest grossly impaired learning in school. Academic performance is, of course, influenced by intellectual potential, familial aspirations, the child's own motivation, and the adequacy of his schooling. School performance is obviously multidetermined. However, in the presence of adequate intelligence, a reasonably normal family environment, and grossly adequate teaching, the single most common cause of school underachievement is minimal brain dysfunction. The difficulties, as noted, are most manifest in reading, spelling, and arithmetic. Difficulties stem from either attentional or cognitive deficits, or both. Management, as will be discussed, depends on where the chief problems lie. Attentional deficits may often be largely eliminated through the use of medication; whereas the perceptual-cognitive deficits demand special educational remediation. In the past, MBD children have often been labeled as underachievers and unsuccessful attempts have been made to pinpoint their problems in terms of familial conflict and neurosis. MBD may persist well into adolescence and continue as a source of underachievement. One study of adolescent underachievers revealed that two thirds of the nonretarded children had had typical signs of MBD in earlier childhood (6).

Family Problems

Family problems are an extremely common concomitant of MBD. They exist for two reasons. First, as will be mentioned later, many of the parents have had, and continue to manifest, signs of MBD themselves and thus show antecedent problems themselves. For this reason, they are a group constituted to least well tolerate and to overreact to the problems produced by MBD children. Secondly, the child's behavioral difficulties engender and increase normal family frictions. The MBD child's refractoriness to discipline and his difficulty in interpersonal relations cause problems with parents and siblings. Disagreements between parents regarding discipline and child rearing techniques are frequent, and it is all too common for each to attribute responsibility for the child's problems to the other. Unfortunately, these difficulties are often exaggerated iatrogenically when the physician, psychologist, or social worker attributes the child's difficulty to his parents' failings and serves to aggravate an already difficult situation. The diagnostic implication is clear but frequently missed: one should clearly explore a child's developmental, temperamental, and cognitive characteristics even if he comes from a disturbed background which might serve to explain his present level of functioning.

Neurological Signs

A variable fraction (again depending upon the source of referral), perhaps one half, of MBD children demonstrate soft neurological signs. The term "soft" is used in three senses: (*a*) to designate behaviors (sometimes perceptual or cognitive) which are deviant because of persistence beyond a normal age (e.g. right-left confusion is normal for a four year old but not an eight year old) (*b*) to designate signs which are either inconsistently present or for which poor inter-rater reliability exists (although, interestingly, higher reliability may exist for such attributes as clumsiness as opposed to hard signs such as reflex changes) (7), and (*c*) to designate signs not correlated with any known underlying neurological lesion. Among the soft signs frequently reported are impaired fine motor coordination, visual-motor difficulties, right-left confusion, impaired rapid successive movements, generalized clumsiness, poorly articulated speech, etc. It should be reemphasized that the presence of such signs may strengthen the diagnosis but that their absence in no way rules it out.

Repeated electroencephalographic studies of MBD children have been done over the years. In general, they show a slightly increased prevalence of a variety of abnormalities among MBD children but these are of statistical, not practical, significance (8, 9); a large number of children meeting all the other criteria for MBD fail to demonstrate EEG deviances.

Congenital Anatomical Stigmata

Recently, a number of investigators have reported an increased frequency of minor anatomical abnormalities in children with MBD (10, 11). These abnormalities were first described in schizophrenic children and overlapped those seen in Down's syndrome: epicanthal folds; hypertelorism; low set and malformed ears; high, arched palate; short, incurving fifth finger; single palmar crease; abnormally long and webbed third toe; strabismus; and, perhaps, skull circumferences which are outside the range of normal limits. Recently, these abnormalities have been reported to be correlated with a history of hyperactivity in the MBD child's father or a history of first trimester bleeding during the mother's pregnancy with the affected child (11).

Psychological Test Performance

There are no pathognomonic psychological test characteristics of MBD children. Population surveys indicate that their intelligence tends to be normal or only slightly below (12). Children having the perceptual-cognitive abnormalities described generally manifest poor performance on tests specifically designed to tap these areas of functioning. (Such tests are those used primarily by special educators and not by clinical psychologists.) As with the case of the neurological signs, absence of psychological test abnormalities in no way rules out the diagnosis.

PREVALENCE

As is obvious, the boundaries of MBD are fairly hazy, and accordingly it would seem likely that prevalence estimates of the disorder would vary considerably. Nonetheless, and despite the fact that different investigators have employed different diagnostic criteria, studies in different geographical areas have come up with a surprisingly constant prevalence of MBD which appears to be of the order of magnitude of 5 or 10% (13–15). The disorder, like most psychiatric disorders of childhood, is much more common in boys, with sex ratios ranging from two to one and higher. A figure of 5–10% may appear large, but may not seem so in view of the fraction of adults who suffer severe psychiatric disability (estimated to be on the order of 15% in one psychiatric epidemiological study) (16). As mentioned, in clinics treating psychiatrically disturbed, primary-school age children, MBD, broadly construed, probably constitutes at least 50% of the cases referred.

DIAGNOSIS

As in the case of most psychiatric disorders, the diagnosis of MBD is principally made by history. Psychological, neurological, and anatomical abnormalities may be corroborative, but absence of such findings in no way rules out the presence of the syndrome. Historical material should be obtained both from the child's family and school. Teachers and counselors are in an unusually advantageous position to assess the child's behavior since, unlike families, they have been able to establish age norms and have the opportunity to observe directly the child's academic performance and social relations with and in comparison to his peers. Psychiatric histories of children should be structured; failure to inquire directly and specifically produces considerable false negative information. This has been clearly documented by one study (17) in which it was found that in less than half of the instances did the parents spontaneously report the presence of such seriously deviant behavior as stealing, destructiveness, chronic bullying, serious lying, and truancy. If parents fail to report such signs, it can be anticipated that they will fail to report an even higher fraction of less deviant behaviors such as restlessness, short attention span, etc. If learning problems are present, the physician, after having ruled out such factors as malnutrition, sensory impairment, grossly deviant family background, and inadequate schooling, should obtain measures of the child's intelligence and standardized achievement scores. If the achievement measures are discrepant with intelligence, i.e. lower than would be predicted, special educational testing is indicated. As mentioned, routine psychological tests are generally not helpful.

Examination of the child himself is usually of little value. When the observations of teachers, social workers, parents, and physicians are intercorrelated, the single observer whose observations fail to correlate with the others is the physician. In one study, approximately 75% of children referred to physicians for hyperactivity failed to display such behavior in the physician's office (18). This finding is consistent with the frequent observation of teachers that the child who is ordinarily disruptive may be self-contained in a one-to-one situation.

The two major differential diagnoses are social or reactive difficulties and borderline schizophrenia. With regard to the former, it cannot be overemphasized that disruption in the familial environment in no way rules out the concomitant presence of MBD; since the disorder is often genetic, it is only too common to find that a child has both MBD and a "bad" family. MBD signs and symptoms are frequently aggravated by difficult family circumstances but only very rarely caused by them.

Examination by the physician is critical when the possibility of borderline schizophrenia exists. Some borderline schizophrenic children manifest MBD characteristics such as inattentiveness and emotional lability. Borderline schizophrenia is suggested by marked fearfulness, pronounced phobic symptomatology, undue preoccupation with violence or sexual matters, and chronic avoidance of social contact. The differential diagnosis is frequently difficult to make.

There is no critical diagnostic test for MBD. Since absolute criteria have not been established, historical, neurological, and psychological data can only suggest the diagnosis. Given this state of affairs, one must decide to use either loose or stringent criteria and accordingly to over- or under-diagnose. As will be mentioned in the section on management, it is probably most useful to err on the side of overdiagnosis, since doing so will suggest the initiation of a therapeutic drug trial. Since drug therapy is unusually safe and the latency until determination of efficacy is short, it is to the child's and family's advantage to employ such a trial. By erring in this direction, one may treat a few children for whom drug management was unnecessary, but one will treat many more children in whom the diagnosis might have been missed.

ETIOLOGY

Probably the earliest descriptions of the MBD syndrome were in children who had developed behavioral abnormalities following, and therefore attributed to, injury or infection of the brain. Following the World War I pandemic of von Economo's encephalitis, reports appeared describing typical and severe MBD behavioral sequelae of that illness in children (19, 20). Later reports described similar behavior patterns occurring in association with a number of suspected or documented causes of neurological damage including head injury, pertussis, lead poisoning, pre- and paranatal medical complications, and prematurity (21–24). It is from these early etiological associations that terms such as "minimal brain damage" and "minimal cerebral damage" stemmed. There are two sources of evidence which suggest that MBD may arise from congenital, not traumatic, causes. The first is the increased prevalence of anatomical stigmata, which suggests either fetal maldevelopment or a genetically transmitted disorder. The second source of evidence comes from repeated observations of clustering of MBD disorders in families. Clinicians have noted a seeming increase in the childhood history of MBD in parents and of other psychiatric illness in parents and siblings. Dyslexia has been shown to exhibit such a family clustering (25, 26) (and dyslexia may be synonymous with the perceptual-cognitive difficulties seen in some MBD children). Interestingly enough, one study showed an increase of MBD-like behavioral disorders among the siblings of dyslexics, suggesting a linkage between the behavioral and cognitive manifestations of the supposed syndrome. Familial clustering of psychological disorders is a finding ambiguous to interpret. It might represent either psychological or social transmission or genetic transmission. One technique for separating the effects of nature and nurture is to study adopted children. In such instances, the parents supplying the genetic loading and those supplying the environmental circumstances are separated, and accordingly one may determine which group of parents is more closely related to MBD behavior. Safer (27) conducted a study of the behavioral characteristics of the sibs and half-sibs of MBD children raised in foster homes. He found that the syndrome occurred with greater frequency among the sibs than the half-sibs, and that in both groups it occurred with greater frequency than in the general population (although the sibs and half-sibs had never been exposed to the same psychological environment that the proband, the child raised with the biological parents, had. Morrison & Stewart (28) have studied the frequency of psychiatric disorder among the biological and adopted parents of hyperactive children. They found an increased frequency of alcoholism, hysteria, and a history of hyperactivity in childhood in the biological parents of hyperactive children as compared to the adoptive parents of hyperactive children and the biological parents of normals. These findings, too, point to a genetic component. The increase of non-MBD diagnoses among the biological parents of MBD children may at first seem surprising. All that may be said is that psychiatric disorders do not seem to "breed true" and that the finding of increased frequency of other diagnoses among the relatives of psychiatric patients is recurrent (29, 30). It is the impression of some clinicians that the genetic form of the disorder is more common than the traumatic form, but hard data are not available.

With regard to the intimate patholophysiological mechanism by which genetic, and presumably biochemical, abnormalities are transformed into the behavioral abnormalities observed, there is no information. This author has suggested elsewhere (57, 58) that genetic transmission together with the fact that many of these children respond dramatically to treatment with stimulant drugs (methylphenidate and the amphetamines) and the tricyclic antidepressants strongly suggests the possibility of an inborn error of monoamine metabolism: specifically, a decrease in levels of serotonin, dopamine, or norepinephrine or in activity of systems in which they are neurotransmitters. [This not strikingly original hypothesis is based on an argument similar to that which has generated the monoamine theory of depression (33).] Some animal models exist. Certain strains of dogs (34) and rats (35) show certain behaviors analogous to MBD children and show similar responsiveness to stimulant medications. Recently, it has been shown that rat pups nursed with mothers who had been fed lead manifested hyperactivity. This hyperactivity was accompanied by a decrease in brain dopamine. Whether this hyperactivity in rats is responsive to stimulant drugs has not yet been demonstrated (36). These results are of interest because of the decreased dopamine (which had been hypothesized as a possible cause of MBD) and because of a reported increase in subclinical lead poisoning in hyperactive children (37).

PROGNOSIS

In the past, it had been generally believed that the MBD syndrome was a benign psychiatric illness with a good prognosis: ostensibly, children outgrew it at puberty or early adolescence. This belief was conveyed by the synonym (not yet mentioned) "developmental disability," which suggested that the disorder was self-limited. Clean, neat prognostic data are not yet available. The reason is twofold. Since the syndrome has attracted attention only comparatively recently, investigators are forced either to examine old records, identify children who would now be diagnosed as having MBD, and then follow them up (with the inaccuracies in identification that this entails), or to perform retrospective studies in which adult psychiatric populations are examined for childhood histories of MBD (with the accompanying sampling biases this introduces together with the inaccuracies of retrospection). The available anterospective follow-up studies are necessarily relatively short-term. A further compounding difficulty is the probable etiological heterogeneity of the syndrome. By analogy, we would anticipate ambiguous data following up patients with dropsy if these included individuals with congestive heart failure, renal disease, perpheral venous disease, and cirrhosis. Retrospective longitudinal studies show an increased prevalence of sociopathic disorders and psychosis among severely impaired MBD children (38–40). However, it must be emphasized that many of these children may have had maximal rather than minimal brain dysfunction and that it is possible, with current diagnostic techniques, that some of these later-to-become schizophrenic individuals would have been diagnosed as borderline schizophrenics at the time. Menkes' study (38) reported anecdotally that some of the adults followed up continued to have attentional difficulties into the third decade (a report compatible with that of the parents of some MBD children). Mendelson (41), following a cohort of hyperactive children anterospectively for two to five years, found that approximately one half of the children did indeed improve moderately, that one quarter remained unchanged, and that approximately one quarter manifested some degree of improvement. Overall, the symptoms of restlessness, distractibility, impulsiveness, excitability, and aggressiveness seemed to persist in most of the children, continuing to be associated with poor performance in school and low self-esteem; in addition nearly 60% had had some contact with the police and 18% had been seen at juvenile court. These figures constitute a much greater risk for "acting out" activities than is to be expected in the general adolescent population. Another anterospective study (42) following children from the initial ages of six to twelve for a period of five years likewise tended to indicate continued, unabated MBD symptomatology in approximately one third of the group. Attempts to determine reliable prognostic variables has, of yet, been relatively unsuccessful. An obvious source of difficulty is that of dividing children into meaningful homogeneous subgroups.

A number of retrospective studies of adult psychiatric patients had revealed an increased prevalence of MBD history and signs in childhood (43, 44) and concurrent neurological abnormalities (45). These characteristics have been found in a variety of diagnostic subtypes but apparently predominantly among young adults with personality disorders characterized by impulsivity, agressiveness, and emotional lability, who are apt to receive such diagnoses as sociopathic and emotionally unstable personalities. A study mentioned before of a clinically less disturbed but very common group, adolescent underachievers (46), found that approximately two thirds of a successively admitted sample of nonretarded, underachieving children had had clear-cut signs of the MBD syndrome in early childhood.

The effects of familial and social experience and treatment on the outcome of MBD is unclear. One author (47) found that premature infants (a group at higher risk for MBD) were three times as likely to show maladjusted classroom behavior when exposed to medical and psychosocial stress. The Minde et al study (48) and clinical experience suggests that adequacy in coping ability of the family may in some instances influence the prognosis.

MANAGEMENT

Treatment of the child with MBD varies from case to case, depending on which symptoms are most salient. Since the pathogenic mechanisms are unknown, treatment is necessarily empirical and symptomatic. Nonetheless, a large fraction of children with MBD can be greatly benefited, at least on a short-term basis, by relatively simple intervention. The four principal therapeutic modalities are: family education and counseling, medication, remedial education for the child, and psychotherapy for the child (2).

In working with the family, the first step is an explanation of the problem. Most parents are apt to attribute their MBD child's difficulties either to their own failings or to some sort of inexplicable malevolence on the part of the child. It is difficult, but necessary, to explain to the parents that behavioral abnormalities may be a manifestation of abnormal functioning of the brain. (A comparable difficulty is frequently encountered with adult endogenous depressives who interpret their problems as psychological in origin and are reluctant to receive drug treatment.) The existence of physiologically caused problems in a child obviously does not rule out the coexistence of aggravating familial problems, and the physician must always be careful to allow himself operating room in dealing with such problems, i.e. by not commiting the family to the view that they are entirely free of responsibility. A second important feature of management, too often neglected, is to inform the parents of the prognosis and necessary treatment plan. It is particularly important to make the parents realize that the disorder is often chronic and though it may diminish at or around the time of puberty, it may continue well into adolescence or beyond. They must fully understand that treatment, to be effective, may well have to be continued over a period of many years. Too often, particularly when medication is immediately effective, parents are lulled into a false sense of security and stop treatment, although persistent (perhaps less severe) problems may continue for many years. Most parents (and many patients) have a short-term model of therapy (e.g. penicillin for middle ear infection) and if not well informed will often stop treatment after a relatively short period of time. A third point, important to communicate, is that in many instances a variety of techniques will have to be employed in varying combinations and that there may be no simple (one-shot) remedy for the child's difficulties.

The most effective single treatment for MBD children is medication. Responsiveness of many of the signs to the stimulant drugs is so prompt and marked as to have suggested their use as a diagnostic test. The amphetamines were first employed in the treatment of hyperactive children by Bradley in 1937 (49). He initially described their results as paradoxical in that they seemed to produce effects antithetical to those produced in adults. By this he referred primarily to their decreasing activity level and their failure to produce euphoria. In fact, it would seem that some of the effects of the stimulant drugs are similar in both children and adults while some are different. In the former category are insomnia, anorexia, and increased levels of activation as measured by psychophysiological measures (primarily skin resistance or conductance). Nonetheless, the clinician is struck by the paradoxical effects which are indeed very different from those produced in adults. Stimulant drug treatment is, when effective, one of the most strikingly effective physical treatments in psychiatry, comparable, perhaps, only to electroconvulsive therapy in the treatment of serious endogenous depression. The slowing down of the child is the least remarkable effect; many reversible chemical poisons can slow children down. Stimulants and tricyclic antidepressants have not only a quieting effect, but also, and of greater importance, a characteristic of considerably greater interest: an effect on complex psychological functioning. While many reversible chemical poisonings are capable of quieting, often at the same time impairing consciousness and higher mental functioning, the effect of stimulant and antidepressant drugs on MBD children seems to be radically different. During the duration of their action, they often produce immediate psychological growth. In addition to becoming calmer and displaying a greater attentional span, children may display more mature cognitive, social, and interpersonal functions than they have ever shown. Previously unable to think ahead, they do so; previously unresponsive to social "shoulds" and "should nots," they become responsive; previously unaware of their effect on peers and parents, they transiently develop this ability. This behavioral reactivity, in addition to being of considerable clinical benefit, suggests that these drugs are (as commented on briefly in the section on etiology) operating at a point relatively close to the origin of the etiological chain.

A number of well-controlled studies have demonstrated that between two thirds and four fifths of MBD children show a favorable response to the stimulant medications when they are used properly (50–52). The drugs most frequently employed are the amphetamines and methylphenidate. D-amphetamine or DL-amphetamine is used in a dose range of 5–50 mg/day, and methylphenidate is administered in a dose range of 5–120 mg/day. Dosage must be carefully individualized and treatment responsiveness based on reports from teachers as well as parents. The basic principle of administration is to use these drugs as one would digitalis, increasing the dosage until an optimum treatment response is reached or side effects become objectionable. In most instances the therapeutic ratio is acceptable. In some instances one must sacrifice good behavioral results in order to minimize unpleasant side effects. The latter consist mainly of anorexia, insomnia, and increased nervousness. Allergic or acute toxic reactions are extremely rare. It is for this reason that therapeutic trials of medication are indicated in instances where the diagnosis of MBD is uncertain. The probability of benefit is not offset by any appreciable risk. No chronic toxic reactions have been demonstrated over relatively brief (several years) periods of time although toxic reactions with the amphetamines have been reported in amphetamine abusers who employ the drug in dosages several hundred times as high. Nonetheless, as is the case in chronic administration of any drug, one must be wary of subtle accumulative toxicity. In particular, one must suspect cardiovascular side effects although none has been reported. No abuse of stimulant drugs has yet been reported in older individuals who received them in childhood for the treatment of MBD.

A number of reports (53–55) indicate that the tricyclic antidepressants are often of benefit in MBD children; in that these drugs are nonabusable and have a 24-hr duration of action, they might seem to be the drugs of choice. Unfortunately, the clinical impression is that even when effective they are not as effective as the stimulant drugs and that, unlike the stimulant drugs, children often become tolerant to them. In addition, no long-term, chronic toxological studies have been performed in either children or adults. Another group of drugs that has been used by some clinicians is the family of antipsychotic drugs. These, too, are nonabusable but recent evidence would indicate that they diminish attention span and hence learning ability (56) and clinical experience would indicate that they do not increase responsivity to socialization. In addition, they produce a number of unpleasant side effects and occasional dangerous idiosyncratic reactions; their chronic administration in adults for long periods of time has been reported to produce a permanent neurological syndrome, tardive dyskinesia. Because of these side effects, I restrict their use to those rare children for whom they are absolutely necessary, children who are so destructive or hostile as to be almost completely unmanageable.

In children with learning problems, special forms of remedial education are mandatory. In those children whose underachievement is secondary to behavioral and attentional deficits alone, medication may completely "normalize" the child. In such instances, nonetheless, he has often fallen behind academically and requires "catch-up placement" to acquire the missed material. In children with combined attentional and perceptual-cognitive problems, medication appears to diminish the attentional problems but not to affect the perceptual-cognitive problems. Medication does not seem to be of appreciable benefit to those children whose problems are in the perceptual-cognitive sphere alone. The last two categories of children require the help of special educational teaching methods. Special educational schools of thought and technique abound. Clear documentation for the efficacy of specific methods with specific kinds of children remains to be produced. As always, the existence of conflicting schools raises the question of their relative efficacy since it is usually the case that techniques of intervention which provide clear-cut advantages are rapidly proven. A particular problem confounding the evaluation of such techniques is the fact that in many MBD children learning disabilities disappear without treatment as they grow older. As this is true, the effects of chronic treatment become confounded with those of maturation. In some instances, it seems that otherwise intelligent children are almost totally unresponsive to these educational techniques. In view of the fact that a certain fraction will continue to have serious reading difficulties for the rest of their lives, it is surprising that no one has decided to teach these children as they would blind children (a therapeutic tack that might be suggested by an old phrase for these cognitively handicapped children, "word blind") since many truly blind children are capable of progressing through graduate school. The interaction of medication and special educational techniques, if any, likewise remains to be evaluated.

As previously remarked, there are often problems between the parents themselves and between the parents and the child with MBD. In some instances, these are the consequences of the parents having the same disorder (in somewhat muted form) and in others they are the result of unsuccessful attempts to deal with the child's problems. In many instances, couple or family therapy may mitigate these problems which, although not having caused the problem in the child, often aggravate that problem. In addition, many parents can be benefited by assistance in working out firm, consistent programs of child management (2). In those instances in which medication does not alleviate many of the child's problems, formal programs of behavioral management appear to be useful in ameliorating the more serious behavioral problems. Such programs are indicated by the fact that they can teach seals to play "My Country Tis of Thee" or profoundly retarded individuals methods of self care. Unfortunately, they demand a sometimes exorbitant amount of time and effort on the parents' part and so, in many instances, may not be practical. It seems, also, that even when effective, these techniques may not generalize, i.e. their effects may not manifest themselves outside the home. As in the case of all other management techniques, programs of behavioral management require long-term evaluation. Individual psychotherapy is definitely of very limited value in treating the common symptoms of MBD—in contrast to their good response to medication. However, the child with MBD is not only the victim of his immediate symptomatology but suffers from the indirect consequences produced by the effect of his behavior on others. When his problem has gone untreated for long periods of time he has suffered from repeated academic failure and has met with continuing nonacceptance, censure, and rebuff from parents, teachers, and peers. Such psychological scars are often not relieved, even when medication is extremely effective. In these instances, limited psychotherapy, directed at helping the child to understand the cause and consequences of his problems, together with the teaching of new techniques of adjusting to the world, may be of some benefit.

Management—an Overview

Some hard data exist regarding the short-term efficacy of medication in alleviating the MBD child's problems. No hard data exists regarding the utility of family counseling, family and individual psychotherapy, or remedial educational techniques. The long-term utility of all these techniques is unknown. Regarding long-term treatment, one point bears iteration. In a fairly substantial fraction of MBD children, the difficulties do not disappear at adolescence but appear to persist well into adult life, and, perhaps, in some instances, for the duration of life. It is entirely possible that some individuals will require pharmacological treatment for the rest of their lives. This has been the experience of some clinicians who have continued to administer stimulant medication to MBD individuals well into their 20s and 30s. Clinical observations seem to indicate that such individuals do not become tolerant to the effects of these drugs nor do they become addicted to them because of their euphoric properties. In such individuals, the model of drug treatment would be that of epilepsy or diabetes and not that of depression (a self-limited disorder in most adults). Failure to recognize this state of affairs would lead to spurious pessimism regarding the efficacy of pharmacological treatment. In individuals in whom the problem is life-long, drug treatment during childhood and adolescence might be expected to produce temporary benefits but no long-term benefits since the underlying biological problem would continue to exist following drug discontinuation. Treatment effects in all chronic diseases are difficult to determine and psychiatric illnesses, including MBD, unfortunately constitute no exception.

CONCLUSION

There are several points which should be re-emphasized. First, MBD is an extremely common disorder of primary-school age children and a problem which may persist into adolescence and adult life. Its major signs are chronic behavioral difficulties and academic impairment despite the presence of normal intelligence. In some instances, the disorder does not disappear at adolescence and may persist, perhaps with changed signs and symptoms, into adult life. In its more severe forms, the syndrome may be the precursor of a number of psychiatric disorders of later life including emotional unstable and impulsive personality disorders and sociopathy. The disorder appears to occur frequently on a genetic basis and may be a manifestation of an abnormality in monoamine metabolism. Treatment with stimulant medication or antidepressants is often dramatically effective in ameliorating the behavioral but not the cognitive symptomatology. Finally, it may be the case that a fraction of children afflicted with these disorders do not outgrow them with age, and that such individuals may require drug treatment for indefinite periods of time; in such instances, therapy may be permanently suppressive although not curative.

Literature Cited

1. Laufer, M. D., Denhoff, E. 1957. *J. Pediat.* 50:463–75
2. Wender, P. H. 1971. *Minimal Brain Dysfunction in Children.* New York: Wiley
3. Sprague, R. L., Barnes, K. R., Werry, J. S. 1970. *Am. Orthopsychiat.* 40:615–28
4. Halverson, C. F., Waldrop, M. F. 1973. *Child Develop.* 44:678–81
5. Boder, E. 1973. *Develop. Med. Child Neurol.* 15:663–87
6. Hammar, S. L. 1967. *Pediatrics* 40:373–81
7. Rutter, M., Graham, P., Yule, W. 1970. *A Neuropsychiatric Study in Childhood.* Philadelphia: Lippincott
8. Stevens, J. R., Sachdev, K., Milstein, V. 1968. *Arch. Neurol.* 18:160
9. Wikler, A., Dixon, J. F., Parker, J. B. Jr. 1970. *Am. J. Psychiat.* 127:634
10. Waldrop, M. F., Halverson, C. F. Jr. 1971. In *The Exceptional Infant,* ed. J. Hellmuth, Vol. 2
11. Rapoport, J. L., Quinn, P. O., Lamprecht, F. 1974. *Am. J. Psychiat.* 131:386–90
12. Miller, R. G., Palkes, H. S., Stewart, M. A. 1973. *Child Psychiat. Human Develop.* 4:121–27
13. Huessey, H. R. 1967. *Acta Paedopsychiat. (Basel)* 34:130
14. Winokur, G. 1972. *Comp. Psychiat.* 13:3–8
15. Minde, K. et al 1971. *Except. Child.* 38:215–21
16. Srole, L. et al 1962. *Mental Health in the Metropolis.* New York: McGraw
17. Graham, P., Rutter, M. 1968. *Brit. J. Psychiat.* 114–581
18. Kenney, T. J. et al 1971. *J. Pediat.* 79:618
19. Bond, E. D. 1932. *J. Pediat.* 1:310–14
20. Hohman, L. D. 1922. *Johns Hopkins Hosp. Bull.* 380:372–75
21. Strecker, E. A., Ebaugh, F. 1924. *Arch. Neurol. Psychiat.* 12:443–53
22. Lurie, L. A., Levy, S. 1942. *J. Am. Med. Assoc.* 120:890–94
23. Thurston, D. L., Middelkamp, J. N., Mason, E. 1955. *J. Pediat.* 42: 120–28
24. Knobloch, H., Pasamanick, B. 1966. *Merrill-Palmer Quart. Behav. Develop.* 12:7–26, 27–43
25. Frisk, M., Wegelius, B., Tenhunen, T., Widholm, O., Hortling, H. 1967. *Acta Paediat. Scand.* 56:333–43
26. Hallgren, B. 1950. *Acta Psychiat. Scand.* 65 (Suppl.) 1:287
27. Safer, D. J. 1970. *Behav. Genet.* 3:175–86
28. Morrison, J. R. Stewart, M. A. 1973. *Arch. Gen. Psychiat.* 28:888–91
29. Winokur, G. 1972. *Comp. Psychiat.* 13:3–8
30. Winokur, G., Reich, T., Rimmer, J., Pitts, F. 1970. Alcoholism: III diagnosis and familial psychiatric illness in 259 alcoholic probands. *Arch. Gen. Psychiat.* 23:104–11
31. Boder, E. 1973. *Develop. Med. Child Neurol.* 15:663–87
32. Rutter, M., Graham, P., Yule, W. 1970. *A Neuropsychiatric Study in Childhood.* Philadelphia: Lippincott
33. Schildkraut, J. J., Kety, S. S. 1967. *Science* 156:21–30
34. Corson, S. A., Corson, E. Kuduch, V. 1973. *Proc. Soc. Neurol. Sci.* Presented at Third Annual Meeting
35. Coyle, J. T. Jr., Wender, P. H., Lipsky, A. 1973. *Psychopharmacologia* 31: 25–34
36. Stansberry, K. G., White, R. S. 1974. *Science* 182:1022–24
37. David, O. Clark, J. Kytja, V. 1972. *Lancet* 2:900–3
38. Menkes, M., Rowe, J. S., Menkes, J. H. A. 1967. *Pediatrics* 39:393–99

39. Morris, H. H. Jr., Escoll, P. J., Wexler, R. 1956. *Am. J. Psychiat.* 112:991–97
40. Robins, L. 1966. *Deviant Children Grown Up.* Maryland: Williams & Wilkins
41. Mendelson, W. 1971. *J. Nerv. Ment. Dis.* 153:273–79
42. Minde, K. et al 1971. *Except. Child.* 38:215–21
43. Hartocollis, P. 1968. *Bull. Menninger Clin.* 32:102–14
44. Quitkin, F., Klein, D. F. 1969. *J. Psychiat. Res.* 7:131
45. Hertzig, M. E., Birch, H. 1968. *Arch. Gen. Psychiat.* 19:528–37
46. Hammar, S. L. 1967. *Pediatrics* 40: 373–81
47. Drilien, C. M. 1964. *The Growth and Development of the Prematurely Born Infant.* London: Livingstone
48. Minde, K. et al 1971. *Except. Children* 38:215–21
49. Bradley, C. 1937. *Am. J. Psychiat.* 94:577–85
50. Eisenberg, L., Conners, C. K. 1971. In *Behavioral Science in Pediatric Medicine*, ed. N. Talbot, J. Kagan, L. Eisenberg. Philadelphia: Saunders
51. Steinberg, G. G. et al 1971. *Am. J. Psychiat.* 128:175
52. Millichap, J. G., Fowler, G. W. 1967. *Pediat. Clin. N. Am.* 14:767–77
53. Krakowski, A. J. 1965. *Psychosomatics* 6:355–60
54. Rapoport, J. 1965. *Int. J. Neuropsychiat.* 1:635–42
55. Rapoport, J. 1974. *Arch. Gen. Psychiat.* 30:789–98
56. Sprague, R. 1970. *Am. J. Orthopsychiat.* 40:615
57. Wender, P. H. 1972. *J. Nerv. Ment. Dis.* 155:55–71
58. Wender, P. H. 1974. *Life Sci.* 14:1605–21

Copyright 1975. All rights reserved

DIAGNOSIS AND PREVENTION OF POSTOPERATIVE DEEP-VEIN THROMBOSIS

❖7131

L. P. Le Quesne, D.M., M.Ch., F.R.C.S.(Lon.)
The Department of Surgical Studies, The Middlesex Hospital,
London WIN 8AA, England

INTRODUCTION

Deep vein thrombosis (DVT) is a common and potentially serious condition that may occur after any surgical operation, especially those involving abdominal or pelvic organs, or the hip joint. In the majority of instances the condition is an insignificant, often unrecognized, incident in the postoperative period, but its seriousness lies in the damage it can cause to the venous system of the lower limb, leading in later years to so-called varicose eczema and gravitational ulcers, and more importantly to the risk of the potentially fatal complication of pulmonary embolus. For many years it has been known that there is a discrepancy between the incidence of the condition as revealed by autopsy studies (35–55% as reported by various authors) and the apparent incidence as assessed during life on the basis of physical signs (~2%). Accurate study of the condition during life has been inhibited by the lack of a reliable, readily applicable diagnostic technique.

Some 30 years ago Bauer carried out detailed studies of the condition by means of serial phlebography. Phlebography is still the most accurate method for the detection of thrombi throughout the whole length of the venous drainage system of the lower limb, but for a variety of reasons it is unsuitable for prospective studies of large numbers of patients. However, in recent years several new techniques (ultrasound, impedance plethysmography, the use of ^{125}I-fibrinogen) for the diagnosis of DVT. has been developed. The applicability of these various techniques in clinical practice remains to be determined, but there is no doubt that the use of ^{125}I-fibrinogen for thrombus location has already provided information of great importance about the condition.

THROMBUS LOCATION WITH ^{125}I-FIBRINOGEN

Technique

This method, first established in animals (1) and subsequently applied to humans (2, 3), depends upon the fact that labeled human fibrinogen mixes with the body fibrinogen pool and reacts in the same way as the host fibrinogen. Therefore, when a thrombus forms in the leg the resulting localized concentration of radioactivity can be detected by a suitably collimated scintillation counter placed externally over the limb. To protect the patient the source and preparation of the fibrinogen must be carefully controlled, and to protect the thyroid gland from damage from liberated, free ^{125}I its uptake by the gland must be prevented by the administration of potassium iodide. After the initial injection of the tagged fibrinogen the technique is noninvasive, and involves minimal disturbance to the patient. Estimations can be made at the bedside with a portable apparatus, as often as required, for 7–10 days, by holding the collimator against the limb in positions suitably spaced to cover the whole length of the venous system.

Because of the physical decay of the isotope (half life = 60 days) and the biological decay of the fibrinogen (half-life = ∾ 100 hr), the actual counts over any given point will decline from day to day. To allow for this and enable comparison between results of the examination on consecutive days, at the same time as the legs are

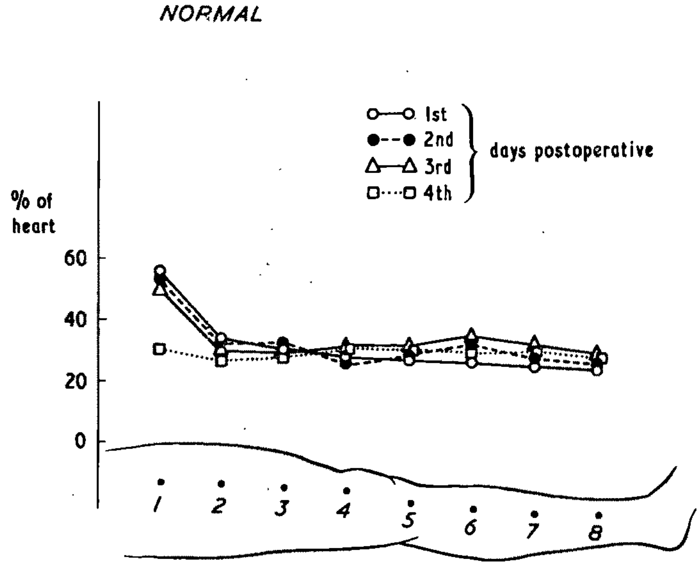


Figure 1 The results of a ^{125}I-fibrinogen study of a normal limb. (Reproduced with permission of the Editor of the *British Journal of Surgery.*)

examined with the scintillation counter a reading is also made over the precordium: the counts at the various points in the leg are then expressed as a percentage of the precordial count, the resultant figure being known as the percentage uptake (4, 5). The findings in a normal limb are shown in Figure 1. It should be noted that the percentage uptake varies very little from day to day and from point to point except with respect to the upper point, at the level of the inguinal ligament: at this level the counts are influenced by the great vessels in the pelvis and free ^{125}I in the bladder, so that the method is unreliable in the detection of thrombi in the upper part of the femoral vein. The method cannot be applied to the detection of thrombi in the iliac veins because of the high background count.

Reliability of Thrombus Location

The presence of a thrombus is indicated by a rise of 20% or more in the percentage uptake at one or more points in the leg compared with the neighboring points (4, 5), the rise being sustained for 24 hr or more (Figure 2). By comparison with phlebography, a number of studies (4, 6) have shown that the ^{125}I-fibrinogen technique is extremely accurate in the detection of thrombi in the calf, popliteal vein, and lower two thirds of the femoral vein. However, it should be noted that the

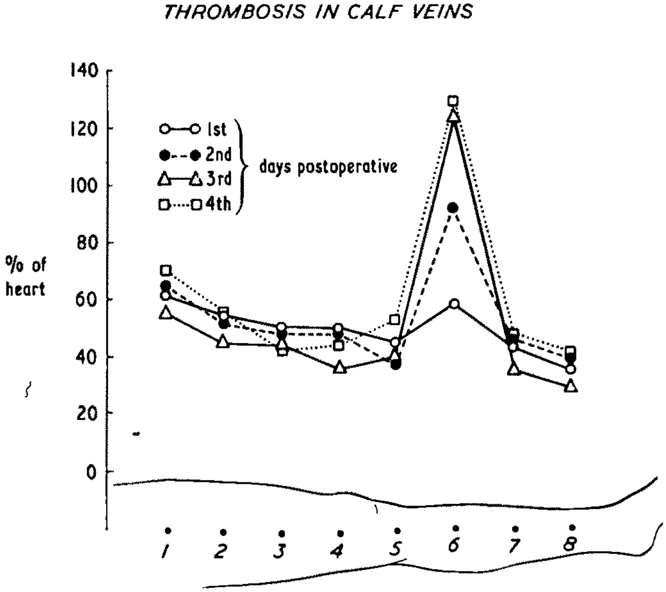


Figure 2 The results of a ^{125}I-fibrinogen study showing the development of a calf vein thrombosis. Note the sustained rise of the percentage uptake over a point in the calf. (Reproduced with permission of the Editor of the *British Journal of Surgery.*)

method is essentially nonspecific in that it detects localized aggregations of fibrin, so that it cannot be applied to patients with incisions in the leg, and care must be taken to exclude superficial thrombophlebitis and other causes of fibrin deposition (e.g. hematoma formation).

The Incidence of Deep-Vein Thrombosis After Operation

Originally used in patients suspected of having DVT, the greatest value of the ^{125}I-fibrinogen test has come from its use in prospective studies. There is now abundant evidence that DVT occurs in some 30% of patients over the age of 40 undergoing major operations (Table 1), and this figure has been confirmed by many subsequent studies. These figures apply to patients of European extraction, and there is evidence from ^{125}I-fibrinogen studies that the condition is less common in other races (7, 8).

Influence of Age and Other Factors

As shown in Figure 3, the incidence of the condition increases from the age of 40 upwards. There is no evidence, by the use of this technique, of the incidence of the condition below the age of 40 yr, and the apparent decline in incidence in the 9th decade is probably misleading and related to the selection of patients suitable for operation at this age. There is no difference in the incidence of the condition in the two sexes, and no convincing evidence that the condition is more common in one leg than the other (I.C. Gordon-Smith, personal communication). Prostatectomy and reconstruction of the hip joint appear to be followed by a high incidence of DVT, but apart from these no other operations carry a particular risk.

Taking into account these various factors, the evidence suggests that DVT occurs in some 30% of all patients over the age of 40 undergoing a major operation, especially those within the abdominal cavity.

Site of Thrombus Formation and Method of Spread

All workers using this technique have found that in > 90% of instances the thrombus is first located in the calf, and in some 90% it remains localized there.

Table 1 Incidence of postoperative DVT, as diagnosed by the ^{125}I-fibrinogen technique

Author	Number of patients[a]	Number of patients with DVT	Percentage of patients with DVT
Flanc et al (6)	65	23	35
Lambie et al (9)	111	49	44
Middlesex Hospital Series (1971)	388	111	28.5
Total	564	183	32.4

[a] All the patients were over 40 years old, and the majority underwent a major operation.

In approximately 30% of patients with DVT the condition is bilateral. Proximal spread occurs in ~ 10% of cases, and may occur in two ways: commonly it occurs as a continuous, centripetal spread of the initial calf thrombus (Figure 4), but occasionally a second, apparently separate thrombus develops in the thigh proximal to the initial calf thrombus. Rarely the initial detected thrombus is in the thigh.

Time of Onset

The majority of thrombi are detected within 48 hr of surgery (Figure 5), and some have even been detected within 2 hr (6). It is, however, important to note that ~ 10% are first detected several days after operation.

Relationship to Physical Signs

It has long been recognized that physical signs are an unreliable guide to the development of DVT. As shown in Table 2, the extent to which this is true is emphasized by studies comparing the incidence of the condition as indicated by clinical examination of the legs and as shown by the ^{125}I-fibrinogen technique. It should be noted that the generally accepted physical signs of tenderness, swelling, Homan's sign, etc not only fail to diagnose the condition in a significant percentage of patients, but also are falsely positive in other patients. The significance of these findings in relation to studies of DVT based upon clinical examination of the legs needs no emphasis, indeed these findings go far toward explaining the discrepancy between the incidence of DVT as found at autopsy and as assessed on clinical grounds.

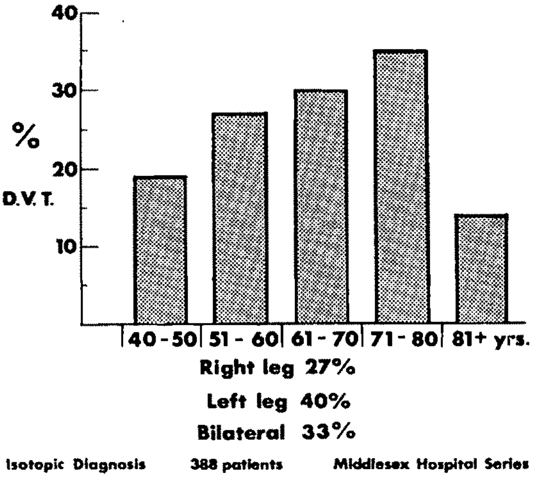


Figure 3 The relationship of age to the development of postoperative DVT.

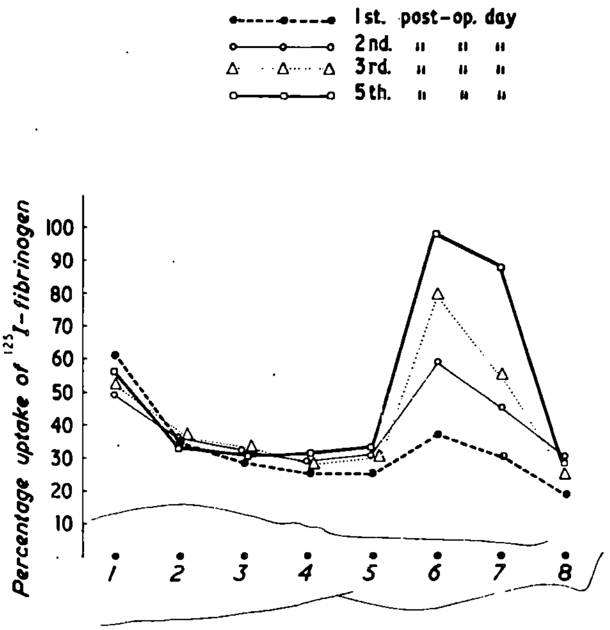


Figure 4 The proximal extension of a calf thrombus.

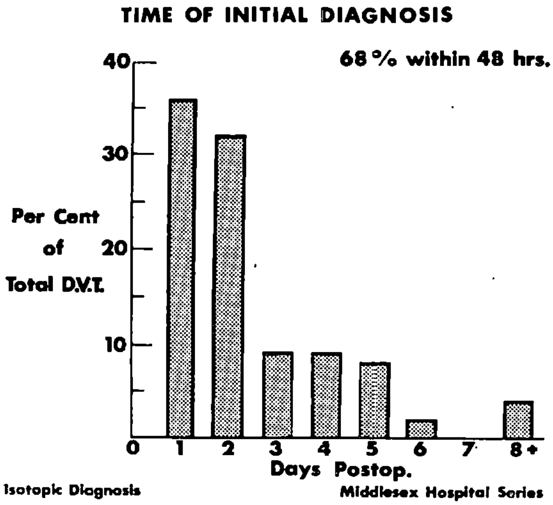


Figure 5 The time of initial diagnosis of DVT in 388 patients studied with ^{125}I-fibrinogen.

Table 2 Accuracy of clinical signs in the diagnosis of DVT (391 patients studied)[a]

Isotopic Diagnosis	Positive in 129 patients: clinical evidence of DVT in only 46 of these.
Clinical Diagnosis	Positive in 70 patients: isotopic diagnosis positive in only 46 of these.

[a] From the figures of Lambie et al (9) and the 1971 Middlesex Hospital Series. Note that clinical signs are present in less than half the patients with isotopically diagnosed DVT, and that in over 30% of those with clinical signs suggesting the presence of DVT the diagnosis was not substantiated by the ^{125}I-fibrinogen examination.

The Prevention of Deep-Vein Thrombosis

Virchow postulated that three factors lead to the development of intravascular thrombosis: stasis, changes in the blood, and changes in the vessel wall. To date there is little factual information about the role of changes in the vein walls in the development of DVT, but the former two factors are thought to play important roles in its pathogenesis after operation. Many attempts have been made to diminish the incidence of postoperative DVT by altering these factors, and in recent years the efficacy of these methods has been studied by the ^{125}I-fibrinogen technique.

The Prevention of Venous Stasis

It is well established that during surgery and for some hours thereafter there is a slowing of the venous return from the lower limb, and this is thought to be an important factor in the causation of DVT. To combat this venous stasis it is conventional to encourage active leg exercises early in the postoperative period, but Flanc, Kakkar & Clarke (10) were unable to show that such exercises had a significant effect in diminishing the incidence of DVT as detected by the ^{125}I-fibrinogen, except possibly to a slight extent in patients over 65 years old.

In contrast to these findings, Browse & Negus (11) showed that intermittent, galvanic stimulation of the calf muscles during surgery resulted in a significant diminution of DVT, as does passive dorsi-plantar flexion of the ankle (12). Similarly, Hills et al (13) showed that intermittent pneumatic compression of the calf during surgery significantly diminished the incidence of DVT, and they also showed that the condition is less easily prevented in patients undergoing an operation for malignant disease than for benign conditions. In all three of these studies DVT was detected by the ^{125}I-fibrinogen technique, and it is to be noted that in these studies the prophylactic treatment, in contrast to postoperative leg exercises, was active during the operation.

The Use of Antithrombotic Agents

Surgery is accompanied and followed by a number of changes in the blood (e.g. increased platelet stickiness, raised levels of Factor VIII, increased fibrinogen level etc) which make it, at least potentially, more liable to clot; these changes are loosely referred to as "the hypercoagulable state." For many years attempts have been made to diminish the incidence of postoperative DVT and pulmonary embolus by the administration of various drugs which impede the blood-clotting mechanism: prominent among these has been the use of long term anticoagulants such as warfarin, and there is persuasive evidence (14, 15) that this diminishes the incidence of pulmonary embolism. Pinto (16) was unable to show that the incidence of isotopically detected DVT was influenced by a similar schedule of warfarin administration, with the loading dose given 1 hr prior to surgery. However, his study provided no evidence about the incidence of pulmonary embolism, and it is certainly possible that warfarin in this dosage, while not diminishing the incidence of calf vein thrombosis, does prevent spread of the thrombus to a size capable of causing serious pulmonary embolism.

By virtue of its rheological properties and capacity to influence platelet function, Dextran 70 is another substance which has been used to prevent DVT, and there is much evidence that it does diminish its incidence in the postoperative period (17, 18, 19).

In a number of publications in the 1960s Sharnoff and his colleagues (20–22) claimed that subcutaneous heparin given in small doses diminished the incidence of postoperative pulmonary embolism. There is new, clear evidence that heparin given subcutaneously in a dosage of 5000 IU q 12 hr for 5–7 days after surgery, with the initial dose given 1 hr prior to surgery, does significantly diminish the incidence of DVT as diagnosed by the ^{125}I-fibrinogen technique (see Table 3). Heparin in this dosage does not cause consistent changes in the conventional tests of blood coagulation, and its efficacy is probably due to its ability to inhibit activated Factor X (23). Given for a shorter period heparin is also effective in diminishing the incidence of DVT in patients with benign conditions (29), but there is evidence that its effectiveness is related to the length of time it is administered after surgery (Gordon-Smith, Le Quesne & Newcombe, 1974; unpublished observations). Given in this dosage heparin does not appear to cause bleeding from the surgical site or at the site of injection, but this point requires further detailed study.

RELEVANCE TO THE PREVENTION OF PULMONARY EMBOLUS

The real importance of postoperative DVT lies in the possibility that it may cause dangerous pulmonary embolism. As emphasized above, the great majority of thrombi detected by the ^{125}I-fibrinogen technique arise in the calf and the majority remain localized there, only $\sim$ 10% spreading higher up the limb. While there is no doubt that calf thrombi can give rise to pulmonary emboli, it is equally certain that the majority of dangerous and fatal emboli arise from higher up in the limb, from the femoro-iliac venous segment. In all the trials mentioned above the efficacy of the various prophylactic regimens was tested by the ^{125}I-fibrinogen technique, and while there is no doubt that they were effective in the prevention of calf vein

Table 3 The published results of the use of subcutaneous, peri-operative heparin in the prevention of DVT[a]

	Controls		Treated Patients	
	Number	Percentage with DVT	Number	Percentage with DVT
Kakkar et al (23)	27	26	26	4
Williams (24)	29	41	27	15
Nicolaides et al (25)	122	24	122	1
Kakkar et al (26)	39	43	172	9.3
Ballard (27)	55	29	55	3.6
Gallus et al (28)	141	21	131	3.8
Middx. Hosp. Series[b]	123	37	80	12.4
Totals	536	30	613	6.5

[a] In all series the initial dose was given 1 hr prior to surgery, and thereafter doses were given q 12 hr for 5–7 days after surgery. In all of these series the heparin produced a statistically significant fall in the incidence of DVT.

[b] This series includes the patients reported by Gordon-Smith et al (29).

thrombosis, the various studies produced no solid evidence that the incidence of pulmonary embolus had been lowered.

Conventionally, DVT is described as starting in the calf and spreading proximally up into the femoral and iliac veins. There is clear isotopic evidence from a number of observers that this form of centripetal spread does occur, and if this is the common method of formation of the dangerous proximal thrombi, the prevention of calf vein thrombosis can be expected to diminish the incidence of pulmonary embolus. However, on the basis of various autopsy and phlebographic studies this view has been challenged, notably by Mavor (30), and it has been suggested that the dangerous proximal thrombi develop independently from calf vein thrombosis and may even be the cause of the distal calf thrombosis. If this is true, then the prevention of calf vein thrombosis may not diminish the incidence of pulmonary embolus, though the prophylactic measure may independently prevent the formation of both distal and proximal thrombi.

From the autopsy evidence (31–33) it is clear that isolated proximal thrombi can occur, but in the majority of instances such thrombi occur in association with calf vein thrombosis. Further, Hunter et al (34) did not find a single example of fresh femoral vein thrombosis unassociated with calf thrombosis. Similarly, in his classical phlebographic studies of postoperative DVT, Bauer (35) found that thrombi originated in the proximal veins in $< 5\%$ of cases. In comparing the information from these various types of study it should be noted that some of the thrombi detected at autopsy may arise as agonal events, whereas phlebographic and isotopic studies give evidence of changes during life. In addition it may be that the development and progression of DVT is different in patients with so-called spontaneous DVT and that associated with the contraceptive pill, than in the condition as it occurs after surgery, and at the present time it would seem wise not to group them all together as a homogeneous condition. Bearing in mind these various considerations, the available evidence indicates that, in the postoperative patient, while proximal thrombi may develop without distal calf vein thrombosis, in the majority of instances the dangerous, proximal thrombi develop in association with and subsequent to calf vein thrombosis, and accordingly methods which diminish the incidence of calf vein thrombosis may be expected to diminish the incidence of pulmonary embolism.

Direct evidence that this is the case is difficult to obtain. The accurate diagnosis of pulmonary embolus during life is notoriously difficult, and at autopsy not only may emboli be easily overlooked (36) but their significance in relation to the cause of death may be difficult to establish. Two studies (37, 38) have reported that the death rate from pulmonary embolus is less in patients receiving peri-operative subcutaneous heparin than in controls. However, both these studies are open to various criticisms, and further evidence is required before a definite conclusion can be reached.

CONCLUSION

The application of the ^{125}I-fibrinogen technique of thrombus location has established a number of basic facts about postoperative DVT: the condition develops in some 30% of patients of European extraction over the age or 40 undergoing major surgery; it most commonly develops in the calf and remains localized there, though in a minority proximal spread occurs; in the majority the condition arises within 48 hr of surgery, and the conventionally described physical signs are a most inaccurate indication of the development of the condition.

It is now clear that there are available a number of methods effective in diminishing the incidence of postoperative DVT, but their relative advantages and disadvantages are not yet clearly defined. There is as yet no conclusive evidence that these various methods diminish the incidence of postoperative pulmonary embolism. It should be noted that the prevention of DVT and pulmonary embolism are not necessarily the same, in the sense that a method of prophylaxis might not diminish the incidence of calf vein thrombosis yet might confine the thrombus to the calf, thus effectively lowering the incidence of significant pulmonary embolus.

Literature Cited

1. Hobbs, J. T., Davies, J. W. L. 1960. Detection of venous thrombosis with ^{125}I-labelled fibrinogen in the rabbit. *Lancet* 2:134
2. Atkins, P., Hawkins, L. A. 1965. Detection of venous thrombosis in the legs. *Lancet* 2:1217
3. Nanson, E., Palko, P., Dick, A., Fedoruk, S. 1965. Early detection of deep venous thrombosis of the legs using ^{131}I tagged human fibrinogen: a clinical study. *Ann. Surg.* 162:438–45
4. Negus, D., Pinto, D. J., Le Quesne, L. P., Brown, N., Chapman, M. 1968. ^{125}I-labelled fibrinogen in the diagnosis of deep vein thrombosis and its correlation with phlebography. *Brit. J. Surg.* 55:835–39
5. Kakkar, V. V., Nicolaides, A. N., Renney, J. T. G., Friend, J. R., Clarke, M. B. 1970. ^{125}I-labelled fibrinogen test adapted for routine screening for deep-vein thrombosis. *Lancet* 1:540–2
6. Flanc, C., Kakkar, V. V., Clarke, M. B. 1968. The detection of venous thrombosis of the legs using ^{125}I-labelled fibrinogen. *Brit. J. Surg.* 55:742–47
7. Hassan, M. A., Rahman, E. A., Rahman, I. A. 1973. Post-operative deep vein thrombosis in Sudanese patients. *Brit. Med. J.* 1:515–19
8. Cunningham, I. G. E., Yong, N. K. 1974. The incidence of post-operative deep vein thrombosis in Malaysia. *Brit. J. Surg.* 61:482–83
9. Lambie, J. M. et al. 1970. Diagnostic accuracy in venous thrombosis. *Brit. Med. J.* 2:142–45
10. Flanc, C., Kakkar, V. V., Clarke, M. B. 1969. Post-operative deep-vein thrombosis. Effect of intensive prophylaxis. *Lancet* 1:477–78
11. Browse, N. L., Negus, D. 1970. Prevention of post-operative leg vein thrombosis by electrical muscle stimulation. An evaluation with ^{125}I-labelled fibrinogen. *Brit. Med. J.* 2:615–18
12. Sabri, S., Roberts, V. C., Cotton, L. T. 1971. Prevention of early deep vein thrombosis by passive exercise of leg during surgery. *Brit. Med. J.* 3:82–83
13. Hills, N. H., Pflug, J. J., Jeyasingh, K., Boardmann, L., Calnan, J. S. 1972. Prevention of deep vein thrombosis by intermittent pneumatic compression of calf. *Brit. Med. J.* 1:131–35
14. Salzman, E. W., Harris, W. H., De Sanctis, R. W. 1966. Anticoagulation for prevention of thromboembolism following fractures of the hip. *N. Engl. J. Med.* 275:122–30
15. Sevitt, S., Gallagher, N. G. 1959. Prevention of venous thrombosis and pulmonary embolism in injured patients. *Lancet* 2:981–89
16. Pinto, D. J. 1970. Controlled trial of an anticoagulant (Warfarin Sodium) in the prevention of venous thrombosis following hip surgery. *Brit. J. Surg.* 57:349–52
17. Bonnar, J., Walsh, J. 1972. Prevention of thrombosis after pelvic surgery by British Dextran 70. *Lancet* 1:614–16
18. Carter, A. E., Eban, R. 1973. The prevention of post-operative deep venous thrombosis with Dextran 70. *Brit. J. Surg.* 60:681–83
19. Bygdeman, S. 1969. Prevention and therapy of thromboembolic complications with Dextran. In *Progress in Surgery,* 7:114–35. Basel: Karger
20. Sharnoff, J. G., Kass, H. H., Mistica, B. A. 1962. A plan of heparinization of the surgical patient to prevent post-operative thromboembolism. *Surg. Gynecol. Obstet.* 115:75–79
21. Sharnoff, J. G. 1969. Prevention of sudden cardiopulmonary arrest in the perioperative period with prophylactic heparin. *Lancet* 2:292
22. Sharnoff, J. G., De Blasio, G. 1970. Prevention of fatal post-operative thromboembolism by heparin prophylaxis. *Lancet* 2:1006–7
23. Kakkar, V. V. et al. 1971. Low doses of heparin in prevention of deep-vein thrombosis. *Lancet* 2:669–71
24. Williams, H. T. 1971. Prevention of post-operative deep-vein thrombosis with perioperative subcutaneous heparin. *Lancet* 2:950–52
25. Nicolaides, A. N. et al. 1972. Small doses of subcutaneous sodium heparin in preventing deep venous thrombosis after major surgery. *Lancet* 2:890–93
26. Kakkar, V. V. et al. 1972. Efficacy of low doses of heparin in prevention of deep-vein thrombosis after major surgery. *Lancet* 2:101–6
27. Ballard, R. M. et al. 1973. Low doses of subcutaneous heparin in the prevention of deep vein thrombosis after gynaecological surgery. *J. Obstet. Gynaecol. Brit. Commonw.* 80:469–72
28. Gallus, A. S. et al. 1973. Small subcutaneous doses of heparin in prevention of venous thrombosis. *N. Engl. J. Med.* 288:545–51

29. Gordon-Smith, I. C., Grundy, D. J., Le Quesne, L. P., Newcombe, J. F., Bramble, F. J. 1972. Controlled trial of two regimens of subcutaneous heparin in prevention of post-operative deep-vein thrombosis. *Lancet* 1:1133–35
30. Mavor, G. E., Galloway, J. M. D. 1969. Iliofemoral venous thrombosis. Pathological considerations and surgical management. *Brit. J. Surg.* 56:45–59
31. McLachlin, J., Paterson, J. C. 1951. Some basic observations on venous thrombosis and pulmonary embolism. *Surg. Gynecol. Obstet.* 93:1–7
32. Gibbs, N. M. 1957. Venous thrombosis of the lower limbs, with particular reference to bed rest. *Brit. J. Surg.* 45:209–36
33. Sevitt, S., Gallagher, N. 1960 Venous thrombosis and pulmonary embolism. A clinico-pathological study in injured and burned patients. *Brit. J. Surg.* 48:475–89
34. Hunter, W. C., Krygier, J. J., Kennedy, J. C., Sneeden, V. D. 1945. Etiology and prevention of thrombosis of the deep leg veins. *Surgery* 17: 178–90
35. Bauer, G. 1940. A venographic study of thromboembolic problems. *Act. Chir. Scand. Suppl.* 61:1–75
36. Morrell, M. T., Dunnill, M. S. 1968. The postmortem incidence of pulmonary embolism in a hospital population. *Brit. J. Surg.* 55:347–52
37. Kakkar, V. V., Corrigan, T., Fossard, D. P. 1973. Prevention of fatal pulmonary embolism using low doses of subcutaneous heparin. *Eur. Surg. Res.* 5:Suppl.2
38. Sagar, S. 1974. Heparin prophylaxis against fatal post-operative pulmonary embolism. *Brit. Med. J.* 1:153–55

Copyright 1975. All rights reserved

TARDIVE DYSKINESIA

❖7132

George W. Paulson, M.D.
Division of Neurology, Department of Medicine, The Ohio State University, and
Riverside Hospital, Columbus, Ohio 43210

For a decade it has been apparent that a peculiar dyskinetic syndrome can occur in patients who receive phenothiazines for prolonged periods of time. Though the syndrome of tardive dyskinesia (TD) was initially controversial, it is now considered a well defined neurologic entity and a major iatrogenic disease. Patients may have irregular or rhythmical movements of the lips, side-to-side shifts of the chin and tongue, or jerky bodily movements in addition to the more classic "bucco-lingual dyskinesia." The trunk can be hyperextended and there may be up-and-down movements of the toes or "piano-playing" restlessness of the fingers.

Tardive dyskinesia can be easily overlooked by the uninitiated. The movements appear voluntary and are inhibited so easily that they seem inconsequential. Although initially considered a suitable puzzle for a few eccentric neurologists to distinguish from Huntington's disease, or an innocuous curiosity for psychiatrists who assessed restless schizophrenics, it is now agreed that all physicians should recognize TD, especially physicians in family practice, public health, or internal medicine.

HISTORICAL FEATURES

At the first major US workshop on TD in 1967 in St. Louis, the clinicians disagreed, and several physicians considered the movements to be inconsequential, not drug-related, and certainly rare. Distinction between these movements and the "stereotypies" of schizophrenia was considered difficult or impossible. No clear clinical description was accepted at that meeting, and the etiology and significance of TD remained a source of quiet, but at times acrimonious, private debate. Since the failure of that initial group to reach agreement, there have been hundreds of discussions in both the European and American literature. In actual fact, TD had been described in detail even before 1967 (1).

The slow acceptance of the phenomenon of tardive dyskinesia is reminiscent of the historical delay four decades ago in linking the late development of parkinsonian oculogyric crises to encephalitis lethargica. The delayed or "tardive" neurologic effects of encephalitis lethargica appeared years after the episode of encephalitis, and as with TD, the postencephalitic phenomena were variable and similarly influenced by psychological factors. One reason for the delay in recognition and acceptance of TD was that the affected population group was from psychiatric patients in state institutions, the hospitalized group which receives the least intensive medical care in America. Patients with TD were often psychotic and didn't complain, and alternatives to the drugs were unpleasant to contemplate. Furthermore, though TD was different in appearance from anything noted before (2), it did overlap in appearance with the stereotypies seen in schizophrenia as well as with the choreas of extrapyramidal disease.

Crane, one of the earliest and most persistent of the investigators of tardive dyskinesia, emphasized that it is often difficult to accept any iatrogenic phenomenon (3). The common rejoiner in an iatrogenic disorder is to assume that the occurrence is inevitable, or alternatively to assert that the severity of the basic illness justified the iatrogenic measures. There is rapidly increasing concern about TD now, enhanced partially because it is apparent that nonpsychotic and noninstitutionalized patients can also be affected with TD (4). TD can even appear in children, and similar disorders can result from drugs other than the phenothiazines.

It is thus seen that patients with TD may appear in the practice of any internist or clinical neurologist, and a few of these patients have been a cause of medical-legal controversy.

CLINICAL FEATURES

TD is identified by sight, and by sight only. There is no specific laboratory examination and no special feature of the neurologic examination which proves the diagnosis. Nevertheless the fact that it can be identified by casual observation does not, apparently, make its recognition any easier. In contrast to most disorders in medicine, the phenomenon is best observed when the examiner doesn't appear to examine. The dyskinesia may be seen while the patient is walking, sitting, or conversing; but may disappear when the patient is alerted. When standing the patient may be hyperextended with a wide-based stance and many patients rock their shoulders. The arms may also extend out from the sides, while the fingers alternately flex and extend. Even when sitting the in-and-out movement of fingers is still observed. Though intermittent and unpredictable the movements appear to be voluntary rather than choreiform. Some patients with TD may also manifest parkinsonian features, but "piano-playing" fingers are not characteristic of parkinsonian tremor and TD doesn't usually produce increased tone. The movements can be inhibited by the patient and may appear for the first time as the examiner leaves. In some resting patients the features of tardive dyskinesia occur incessantly, and because of dystonic or rhythmical movements of the toes and feet, both toes and shoes are traumatized. In some wards of state hospitals, patients are seen with their great toes intermittently peaking through the top of the shoe.

The classic phenomenon of bucco-lingual dyskinesia is more socially incapacitating than gentle toe movements. Approximately 30% of the patients with tardive dyskinesia manifest these mouth, tongue, or lip movements. The tongue may flick in and out of the lips, and this has been referred to as a "fly-catcher movement" although one trusts that the term originated in the days of the encephalitis epidemic when there were more flies to catch. There can be gentle side-to-side chin movements, rhythmical pursing of the lips, intermittent shrugging of the neck and face, or grimacing movements of the lower part of the face. In contrast to Meige's syndrome (5), blepharospasm is not a conspicuous feature of TD, and the extraocular movements in patients with TD are usually intact. The classic mouthing, like chorea, athetosis, or tremor, can be imitated by clinicians, but some patients have writhing movements of the tongue which are absolutely impossible to imitate. Such rippling or writhing tongue movements are seen with the tongue in the floor of the mouth and may disappear when the tongue is extended. In this movement, as in other movements of TD, the phenomenon is best seen when the patient is relaxed and when the physician appears to observe something else, for example inspecting the gums or teeth. Indeed, since a different type of mouthing movement can be due to an edentulous state, examination of the mouth is indicated in all cases of TD.

TD should be clearly distinguished from two other side effects of phenothiazines. The first is the acute dystonic episode that occurs particularly in the young and which is more linked with prochlorperazine and trifluoperazine hydrochloride than with the other neuroleptics. In acute dystonia or "central tetanus," after a few doses the patient develops an acute wave of dystonic posturing of the neck, perhaps with the tongue protruded or the teeth tightly clenched. Anxiety and hypertension can be associated with the muscular tension, and the painful tightness may fluctuate in waves lasting several hours. The phenomenon is so bizarre, and so often is seen when neuroleptics are given for the first time for psychiatric reasons, that the dystonia has been frequently confused with hysteria. Diphenhydramine hydrochloride, caffeine, amytal, i.v. benztropine mesylate, or simple sleep causes the dystonia to melt away. Although benign, the condition can be terrifying to the patient and misleading to the physician. A few patients have required surgical intervention or a tracheotomy and I have seen one patient with dystonia who received emergency craniotomy to relieve a hypothetical and nonexistent increase in intracranial pressure. In that instance the tonic movements were confused with decerebration.

The second and best known neurologic disability resulting from phenothiazines occurs several weeks or months after phenothiazines have been initiated, and is so-called pseudoparkinsonism. By no firm criteria is this parkinsonism "pseudo," although the tremor does appear a bit faster than the tremor of classical parkinsonism. Rigidity and akinesia in drug-induced parkinsonism may be less prominent than in true parkinsonism, but when the patient is up and walking, the classic parkinsonian simian posture and rhythmical "pill-rolling" movements of the fingers may be quite pronounced. The changes in handwriting, skin, and muscle tone in Parkinson's disease can all be seen as secondary effects of the drugs. The patient with drug-induced parkinsonism may have, with or without other parkinsonian features, akathisia or restless, shuffling feet. It is a rare psychiatric ward that doesn't have at least one patient standing slightly off by himself, quietly and rhythmically raising and lowering the feet. The parkinsonism in psychiatric hospitals is universal, so obvious that medical students confuse parkinsonism with schizophrenia, but the patients with tardive dyskinesia are equally obvious to an experienced observer.

What significance does the parkinsonism, akathisia, or even TD have for the patient? The patients rarely complain about the movements, though they may complain about slowing due to a drug-induced akinesia. A similar lack of concern, or frank denial, can be seen in patients with the dyskinesia due to L-dopa, in the stereotypies observed in the congenitally blind, and in the chorea of Huntington's disease. The families of patients with TD may also ignore the dyskinesias or assume the movements are part of psychiatric illness. For some, however, TD is quite incapacitating, and is particularly distressing to nonpsychotic patients whose vanity has not been blunted by institutionalization. Apathy or lack of complaint about the dyskinesia doesn't imply approval even in institutionalized patients, and when asked, some patients do complain of their physical appearance or of the mental changes they associate with phenothiazines.

There may be a psychic corollary to the lack of concern about dyskinesia in that patients complain of mental slowness in parkinsonism and note obtundation when first placed on phenothiazines, but patients often don't complain of either excess motor movements or of incessant flight of ideas. There has been no study of mental processes in patients with tardive dyskinesia, and no evidence so far of a shift in intellectual responses while the movements are occurring. Apparently increased rapidity of association and excess motor responses are not distressing, while sluggishness or a poverty of ideas is unpleasant to most humans.

As with the acute dystonia, a few patients with tardive dyskinesia have severe diaphragmatic, respiratory, or abdominal movements and in these the distress is intense and rapid treatment is mandatory. Patients with TD can have a repetitive cough, forced swallowing, or tongue protrusion to such an extent that feeding or breathing is difficult or the mouth and tongue are painfully traumatized. Some patients with postencephalitic parkinsonism had similar severe incapacity (6).

INCIDENCE

How common is TD, and what population groups are affected? The disorder is seen primarily in older patients and some authors have suggested it is more common in women than in men though Crane disagrees. TD has been linked with organic brain disease, such as lobotomy, dementia, or prolonged electric shock therapy. TD can occur in young patients, and is not rare in private patients who have never been institutionalized at all. There is absolutely no evidence that it occurs particularly in schizophrenics, though this is of course the group which has received the most massive doses of phenothiazines. It is apparent to all authors that total cumulative dosage and duration of therapy with phenothiazines are related to development of TD. Most severe cases have received phenothiazines for longer than six months, often for over three or four years. Chlorpromazine was the standard drug for many patients, but complex mixtures and high doses were frequently employed. Continued use of anticholinergic medications or other drugs to reduce Parkinson's does not ameliorate or prevent TD. It has even been suggested (7) that use of anticholinergics to combat parkinsonism due to phenothiazines *increases* the likelihood of TD. Certainly the use of anticholinergics may tempt the clinician to try higher doses of phenothiazines in the mistaken expectation that side effects are being prevented.

TD is seen in all US mental hospitals, however, and in both Europe and the United States the incidence may be as high as 15% to 30% of patients on the back wards. Only 1% or 2% of chronic patients are severely incapacitated by tardive dyskinesia, but in these the movements can be quite dramatic and distressing either to have or to witness.

There is no known association of TD with other effects of phenothiazines, such as parkinsonism or the tendency of these drugs to produce sudden death. Since phenomena such as keratopathy of lens or increased skin pigmentation also relates to total dosage there is an overlap in these symptoms. Phenothiazines can increase seizure potential, especially in large doses and when first given, but there is no obvious correlation between seizures and the production of TD except that it is unlikely that high doses of phenothiazines would be continued in any patient whose seizures increase while on the medication.

NEUROPHARMACOLOGY OF THE TARDIVE DYSKINESIAS

A recent excellent book by Klawans (7) has reviewed current information on the pharmacology of TD, and the data are too complex to summarize here. Klawans suggests that one theoretical mechanism of TD is a "denervation hypersensitivity." The denervation hypothesis implies that a neuronal system has sustained an anatomic or biochemical insult which permanently enhances its response to neurohumors.

Patients who develop TD have usually received large amounts of phenothiazines, commonly over 1000 gm of chlorpromazine or its equivalent. Often these long-lasting movements first appear when the dosage is *reduced*, which makes this an unusual variety of withdrawal phenomenon.

Klawans, using the denervation hypothesis, postulates that chlorpromazine blocks dopamine receptors, probably at the lipid interface. After prolonged phenothiazine usage the dopamine receptor has compensated for the block and the receptor is very sensitive to the small amounts of dopamine that still remain. Responsivity evolves into a supersensitivity. The blockade or denervation is chemical, may become structural, and finally leads to a secondary hypersensitivity. In a patient in whom TD appears only when the phenothiazines are stopped, hypersensitive receptors responded intensely to normal levels of dopamine after the block was lifted.

Reserpine acts to prevent the storage of dopamine in intraneuronal granules and it may be true that reserpine leads to akinesia and rigidity by means of a reduction in the total stores of brain dopamine. In the case of reserpine a chemical denervation would not be possible at the receptors since the receptors have not been blocked and thus not gradually rendered hypersensitive. L-dopa increases dopamine and can, of course, increase any movements that are secondary to dopamine. L-dopa may act to increase the movements of tardive dyskinesia.

The postulate that anticholinergics not only do not protect but increase the likelihood of tardive dyskinesia suggests an iatrogenic exacerbation of an iatrogenic disease. It is stated that centrally acting anticholinergic agents can worsen tardive dyskinesia in patients. Dyskinesias reminiscent of TD have been seen not only in monkeys and other animals treated with massive doses of L-dopa, but also in animals treated with chronic dosage with chlorpromazine.

PATHOLOGY

The pathological reports of TD have been sketchy. Hunter et al (8) did not find any specific pathological abnormality in the brains of patients who developed irreversible dyskinesias, but one study from the National Institutes of Health suggested that increased glial tissue and deterioration of astrocytic cells in a patient who suffered from tardive dyskinesia was noted. Christensen (9) studied 28 patients with dyskinesia and noted neuronal degeneration in a large number of these patients. This group emphasized the difficulties of pathologic work in any state mental hospital, in view of the numerous complicated diseases, prolonged cachexia, and inadequate chart data usually available.

THERAPY

Prevention is the best, if least dramatic, therapy; but at this time there is no certain way to prevent tardive dyskinesia. Certainly simultaneous use of anticholinergic drugs does not prevent the development of TD. The use of *drug vacation periods* or intermittent dosage has been suggested and for many psychiatric patients days or even weeks off phenothiazines is not hazardous. The concept of drug vacations is not generally recognized, but for medications like methylphenidate, methysergide maleate, and some of the anticonvulsants and other drugs that have cumulative effects, "therapus interruptus" should be considered. It is also clear that the smallest dose that is maximally effective is the optimal therapeutic goal. Phenothiazines in doses higher than 400 mg per day probably do not increase the antipsychotic benefit. In preventing tardive dyskinesia, phenothiazines should be used with special caution in the aged, the very young, and those with brain damage.

Utilization of long-acting phenothiazines and butyrophenones injected monthly in a subcutaneous and slowly absorbable fashion raises alarming possibilities of an effective way to produce tardive dyskinesia, especially in teenagers in whom such a drug regimen is popular.

Once severe tardive dyskinesia is present, there should be a gradual reduction of the phenothiazine dosage. Such reduction can produce a temporary worsening and small amounts of haloperidol, diazepam, or similar drugs may be necessary to control extreme coughing, respiratory distress, or the dyskinetic movements of the mouth. Surgery for TD has been performed by Blaine Nashold at Duke Medical Center, but he reports benefits were slight. The movements are not as permanent as once thought, and at least 50% of patients slowly improve in the six months after drugs are eliminated.

PHILOSOPHICAL ASPECTS AND SOME REMAINING QUESTIONS

Aside from the obvious social and legal aspects of TD, there are many perplexing questions. Why would such a movement be so well accepted by most patients? Is it perceived as normal restlessness or as a tic? Are there any mental changes associated with the movements? Are there other neurologically active agents, such as the anticonvulsants, from which permanent side effects can result? Osteomalacia, hirsutism, or even facial changes have been postulated as due to diphenylhydantoins, for example. Should we begin to plan vacation periods from some of these agents? Why do certain groups seem more vulnerable to TD? In Wilson's disease the younger patient may also have more dystonia and the old one manifest more tremors. Why does the young brain appear more vulnerable to dystonic responses and the older one to parkinsonism or tardive dyskinesia? Are there other agents which affect the psyche or CNS and produce unrecognized but significant side effects? For example, will the child on methylphenidate or dextroamphetamine sulfate, which can rarely produce a tardive dyskinesia, have a changed CNS, leading in adult life to abnormal movements or to psychologic patterns that are not a residue of childhood hyperactivity, but rather secondary to the methylphenidate itself? Can the alcoholic who repeatedly has developed a temporary withdrawal tremor later develop a permanent "benign essential tremor?" In how many other instances does a biochemical lesion eventually lead to a permanent anatomical lesion?

Even more experimentally oriented questions remain. Will the classifications of dyskinesias induced by L-dopa be shown to overlap those of tardive dyskinesia? The EEGs in some patients with tardive dyskinesia seem reduced in voltage, as in some patients with Huntington's disease. Are there EEG or other neurophysiologic peculiarities, even artifacts due to the lingual dyskinesia, that would help prove the diagnosis?

In conclusion: the tardive dyskinesias are a real, disturbing, and important neurologic entity. Iatrogenic in origin, the treatment consists of prevention by caution in use of phenothiazines, and symptomatic management of intractable cases. Awareness by the internist is mandatory, not only to avoid misdiagnosis and legal embarrassment, but also to avoid worsening or creating the condition.

Literature Cited

1. Schmidt, W. R., Jarcho, L. W. 1966. Persistent dyskinesias following phenothiazine therapy. *Arch. Neurol.* 14:369–77
2. Crane, G. E., Gardner, R. Jr., Eds. 1969. *Psychotropic Drugs and Dysfunctions of the Basal Ganglia.* Washington, DC: GPO
3. Crane, G. E. 1973. Clinical psychopharmacology in its 20th year. *Science* 181:124–28
4. Klawans, H. L., Bergen, D., Bruyn, G. W., Paulson, G. W. 1974. Neuroleptic-induced tardive dyskinesias in nonpsychotic patients. *Arch. Neurol.* 30:338–39
5. Paulson, G. W. 1972. Meige's syndrome. *Geriatrics* 27:69–73
6. Jelliffe, S. E. 1929. Oculogyric crises as compulsive phenomena in postencephalitis. Their occurrence, phenomenology and meaning. *J. Nerv. Ment. Dis.* 69:59
7. Klawans, H. L. 1973. *The Pharmacology of Extrapyramidal Movement Disorders,* ed. M. M. Cohen. Basel: Karger
8. Hunter, R., Earl, C. J., Janz, D. 1974. A syndrome of abnormal movements and dementia in leucotomized patients treated with phenothiazines. *J. Neurol. Neurosurg. Psychiat.* 27:219–23
9. Christensen, E., Moller, J. E., Faurbye, A. 1970. Neuropathological investigation of 28 brains from patients with dyskinesia. *Acta Psychiat. Scand.* 46:14–23

Copyright 1975. All rights reserved

THE ROLE OF HORMONE RECEPTORS IN THE ACTION OF ADRENAL STEROIDS

❖7133

David Feldman, M.D.
Department of Medicine, Stanford University School of Medicine,
Stanford, California 94305

Extensive investigation indicates that all steroid hormones appear to have similar subcellular mechanisms of action. This recognition of a general pattern has led to a unified hypothesis of steroid action and has facilitated efforts to elucidate the underlying mechanisms. Although differences may exist in the details of their action, and each steroid possesses unique properties, a generally accepted sequence of events applicable to all steroid hormones has been postulated: (*a*) binding of the hormone to cytoplasmic receptors in target cells and subsequent translocation of this steroid-receptor complex to a nuclear acceptor site, (*b*) activation of the genome by an as yet unknown mechanism leading to increased RNA synthesis and resulting in the induction of new protein, and (*c*) mediation of the biological effects of the hormone by the induced protein. The present discussion will be limited to recent developments involving adrenal steroids and will concentrate on the first phase of the above sequence—the role of the receptor. Reviews of various aspects of steroid action have been published recently (1–5).

THE RECEPTOR SYSTEM

Attempts at isolating and purifying steroid receptors are just now beginning to meet with success and, of necessity, experiments have been performed with crude mixtures of cellular constituents. The hypotheses regarding receptors are therefore, to some extent, based on indirect evidence. Despite this shortcoming, there is a large body of evidence supporting the concept that the binding of steroid to the receptor is necessary for subsequent hormone action (1–5). Although most biological actions of steroid hormones appear to be mediated by a receptor mechanism, some effects, especially at pharmacological doses, may prove to be mediated by separate pathways.

Steroids seem to penetrate all cells easily by simple diffusion although some workers have postulated an entry mechanism in target cells (6). Specific receptors, however, appear to exist only in responsive or target tissue (1–5). The receptors are felt to be cytoplasmic in location and protein in nature with a molecular weight estimated in the range of 50,000 to 150,000. The binding of steroid to receptor is stereo-specific, of high affinity but noncovalent, and of low capacity (saturable) (1–5).

The interaction of an active steroid with a receptor initiates a temperature-dependent transformation or activation of the steroid-receptor complex that permits nuclear translocation to occur (7–11). This activation seems to involve a conformational change in the complex which has been detected by alterations in sedimentation properties (11–13). The depletion of cytoplasmic receptor sites, as well as other evidence, indicates that the steroid-receptor complex migrates to the nucleus rather than the receptor merely transfering the steroid (7–11). The transfer is quite rapid with a half time for cortisol in thymus cells of about 30 sec at 37°C (14).

Within the nucleus, mineralocorticoids pass through an easily extractable form before binding to a nuclear acceptor site on the chromatin which then requires high salt concentrations for extraction (9, 15, 16). The nature of the chromatin acceptor site is unclear, but most authors agree that prior binding to activated cytoplasmic receptor is required to achieve specific nuclear binding. In the hepatoma cell system more glucocorticoid binding sites were found on naked DNA then in isolated nuclei, leading to the suggestion that DNA may be the acceptor and the chromosomal proteins may act to restrict binding to certain regions of the DNA (17). Reconstitution experiments imply that glucocorticoid acceptors may exist in all nuclei and that, if cytoplasmic receptors are provided from appropriate target tissue, nuclear binding of glucocorticoid can be achieved in nontarget tissue (18).

Subsequent to its interaction with the genome, the details of which are as yet unclear, the steroid-receptor complex releases from the chromatin. The factors initiating the off-mechanism are unknown; however, it is felt that the steroid is still unmetabolized at the time of release (19). The receptor appears to recycle back to the cytoplasm with a half time for receptor regeneration of about 30 min (7). Receptor recycling is temperature- and energy-dependent but does not require protein synthesis (7, 8, 11, 16).

MINERALOCORTICOID VERSUS GLUCOCORTICOID RECEPTORS

The adrenal cortex secretes a variety of steroid hormones but the major functional steroids can be divided into two classes: glucocorticoids and mineralocorticoids. The glucocorticoids are represented predominantly by cortisol in humans and corticosterone in rodents, while in all mammals the major mineralocorticoid is aldosterone. Although these hormones differ only slightly in structure, the receptor can discriminate between them. Thus, in the rat kidney, two species of adrenocorticoid receptors have been defined—those receptors putatively assigned a mineralocorticoid function having a high affinity for aldosterone and a lower affinity for corticosterone (20) and those receptors putatively assigned a glucocorticoid function having a high affinity for corticosterone and a lower affinity for aldosterone (21, 22).

These studies suggest that the overlapping actions of corticosteroids (i.e. the mineralocorticoid effects of some glucocorticoids and the glucocorticoid effects of some mineralocorticoids) represent a lack of complete specificity of the receptor for the hormones. The amount of crossover activity is dependent on the relative affinity of the steroid in question for the alternate receptor. For those hormones possessing a low affinity, the concentration of steroid must be comparably increased to result in sufficient receptor site occupancy to initiate physiological effects. Of great importance in this respect is the extent of plasma binding of the steroid since only the unbound (i.e. free) steroid appears to have access to intracellular receptor sites (20). Corticosterone, which might otherwise inappropriately occupy aldosterone receptors because it circulates at roughly a 1000:1 excess over aldosterone, is highly bound in plasma while aldosterone is minimally bound. This differential plasma binding essentially lowers the effective concentration of corticosterone and serves, in addition to the lower affinity for the aldosterone receptor, to prevent corticosterone access to the mineralocorticoid sites under physiological conditions (20).

A third type of adrenal steroid receptor has been demonstrated in the rat kidney. It has a high affinity for corticosterone but unlike the glucocorticoid receptors described above, has low affinity for dexamethasone (23). The function of these receptors has not yet been ascertained.

TARGET ORGAN DISTRIBUTION OF CORTICOSTEROID RECEPTORS

Receptors have been sought in multiple organs in an attempt to establish the tissue sites of direct steroid action. The organ distribution of mineralocorticoid receptors includes the rat kidney (9, 20), parotid (16), duodenal mucosa, liver, spleen, and brain (15). The physiological significance of the relatively low levels of receptor sites in the latter three organs was questioned (15) and, on further examination of the liver, these receptors at least were felt to represent glucocorticoid rather than mineralocorticoid binding sites (24). In terms of glucocorticoid receptors, the usually studied tissues have been rodent thymus (25), fibroblasts (26), liver (27), hepatoma tissue culture cells (28) and, as mentioned above, kidney (21, 22).

The detection of glucocorticoid receptors in new sites has led to interesting speculation on the role of glucocorticoids in regulating additional functions. In the brain, glucocorticoid receptors have been described in the rat limbic system, especially the hippocampus (29); pituitary tumor cells in tissue culture (30); and the retina (31). Although biological correlates for receptor binding in the hippocampus have not been established, steroidal effects on behavior may prove to be related to these receptor sites. The demonstration of receptors in pituitary cells is of great interest because the extent of glucocorticoid binding in these cells was shown to have an excellent negative correlation with ACTH production (30). This demonstration of feedback inhibition related to receptor binding helps to resolve the issue of whether ACTH regulation occurs at the level of the pituitary. The retina has been shown to contain glucocorticoid receptors and the extent of binding has been correlated with the specific induction of the enzyme glutamine synthetase (31). These findings confirm similar results with tyrosine aminotransferase induction in liver (10, 32) and support the contention that receptor binding is related to hormone expression.

Because of the association of glucocorticoid excess and osteoporosis, the recent demonstration of glucocorticoid receptors in fetal rat bone cells is of considerable interest (33). The presence of receptors in bone supports the possibility of direct effects of glucocorticoids at the tissue level and provides a possible mechanism for steroid-induced osteoporosis. In addition, glucocorticoid receptors have now been reported in rat and dog heart (34), rabbit uterus (35), rat and rabbit muscle, testes, stomach and spleen (36), and rabbit lung (37, 38). The generality of glucocorticoid receptors suggests they mediate most glucocorticoid effects and may be more important than previously appreciated in regulating gene expression (36).

RECEPTORS AND DIFFERENTIATION

Of considerable current interest is the role glucocorticoids play in the maturation of immature tissues. Because glucocorticoid administration precociously induces pulmonary surfactant and induces parturition in many species, some investigators have suggested its use to prevent respiratory distress syndrome in the newborn. Glucocorticoid receptors are present in rat (39), rabbit (37, 38), and human (39, 40) fetal lung prior to the appearance of pulmonary surfactant. In fetal rat liver functional receptors also appear to antedate the appearance of glucocorticoid responsivity (41). Some hepatic enzymes which are inducible by glucocorticoid postpartum fail to respond to glucocorticoid in fetal life. To explain the lack of induction at an earlier time in fetal development, it has been suggested that changes in the level of circulating glucocorticoids may determine the onset of glucocorticoid-induced differentiation near term (42). However, since glucocorticoids are circulating at substantial levels and since receptors are present earlier in fetal life, some additional factor appears primary in determining the onset of tissue responsivity in the immature animal (39, 41). This is an example of a situation where the presence of receptors does not necessarily indicate the ability of a tissue to respond to hormones (41).

STEROID-RECEPTOR INTERACTIONS AND CORTICOSTEROID EFFICACY

Tomkins and co-workers (10, 32) have proposed a model that provides a possible explanation for the differences in potency of various steroids. Based on the ability of steroids to induce tyrosine aminotransferase, these workers found they could categorize a wide variety of steroids into four glucocorticoid groups: (*a*) optimal inducers (e.g. cortisol, corticosterone, dexamethasone, and aldosterone), which induce the enzyme to a maximal level; (*b*) suboptimal inducers (e.g. 11-deoxycortisol, 21-deoxycortisol), which induce the enzyme to a submaximal level and competitively inhibit induction by optimal inducers above the submaximal level; (*c*) anti-inducers (e.g. progesterone, 17-α-methyl testosterone), which fail to induce the enzyme but competitively inhibit induction by inducers; and (*d*) inactive steroids (e.g. epicortisol and androstenedione), which neither induce nor inhibit.

The proposed molecular model to explain these findings holds that the cytoplasmic receptor exists in two conformational states, an active and an inactive form. The uncomplexed receptors are predominantly in the inactive form; binding by an inducer steroid converts the receptor to an active form capable of initiating subsequent nuclear events. Suboptimal inducers bind to both forms of the receptor resulting in a submaximal concentration of the active moiety. Anti-inducers bind only to the inactive form of the receptor thereby failing to induce but possessing the capacity to competitively inhibit active steroids (see below). Active steroids thus act as allosteric modifiers of the receptor protein and variations in glucocorticoid potency are the result of gradations in affinity for the receptor and capacity to form an active complex (10, 32).

CORTICOSTEROID ANTAGONISTS

Suboptimal and anti-inducers represent categories of drugs capable of competitively displacing active hormones from their receptor sites thereby antagonizing hormone action. The spirolactones inhibit mineralocorticoids by this mechanism and are probably the most extensively used steroid antagonists. Recent work has explored the nature of the antagonist-receptor complex to determine why it fails to act. Radioactive SC-26304, a very potent spirolactone analogue (43), was found to form receptor complexes in kidney cytoplasm but these complexes failed to translocate to the nucleus and bind to nuclear acceptor sites (12). These data support the earlier findings with progesterone which functions as an antagonist of glucocorticoid action by a similar mechanism (11). Furthermore, sedimentation analysis of the antagonist-receptor complexes has revealed an inability of these steroids to generate the apparent active moiety (11–13). It has been suggested that anti-inducers (antagonists) fail to induce a conformational change in the steroid-receptor complex, thereby failing to expose a nuclear binding site (11).

Another well studied antiglucocorticoid is cortexolone (11-deoxycortisol or compound S) (25). However, the mechanism of antagonism is in doubt in this case because it remains unresolved whether the cortexolone-receptor complex binds within the nucleus (14, 25) or fails to bind (13). If the former theory is correct, an additional mode of action would be required, possibly the failure of the complex to bind to the proper site on the chromatin or its inability to activate subsequent steps in the sequence. Other possible modes of antagonism include agents that diminish the concentration of receptor or that prevent recycling of used receptors to the cytoplasm.

Since binding to the cytoplasmic receptor site is a crucial property of antagonists as well as agonists, competitive binding assays provide an ideal system for screening potential compounds for this purpose (43). Studies of structure-binding relationships can yield data allowing for a rational approach to the design of useful drugs.

CORTICOSTEROID RESISTANCE OF TUMOR CELLS

As stressed throughout this article, current hypotheses on steroid action require initial binding of the steroid to the cytoplasmic receptor in order to obtain a biological response. The absence of receptors therefore indicates an inability of the tissue in question to respond to a given hormone. This concept thus presents a method for testing the receptor hypothesis and, as discussed below, also provides a valuable clinical tool.

Mutant fibroblasts and lymphoid cells resistant to the effects of glucocorticoids have been isolated. These cells, which are no longer sensitive to the cytolytic effects of glucocorticoid, were found to have deficient glucocorticoid binding (26, 44, 45). Recently glucocorticoid receptors have also been observed in human acute lymphoblastic leukemia cells and in some cases of acute myelocytic leukemia (46, 47). Six patients with acute lymphoblastic leukemia who initially had receptors present in their blasts were responsive to combinations of drugs including glucocorticoids. At a time when the patients no longer responded to the drugs, the receptors could no longer be detected in the cells. This led to the conclusion that the appearance of drug resistance was related to a change in receptor content (46). Along similar lines the role of estrogen receptors in the clinical management of breast cancer is discussed elsewhere in this issue (McGuire). It should be emphasized that the presence of receptors in a tissue is not necessarily indicative of the responsivity of the tissue to hormone (41). All of the multiple steps subsequent to receptor-binding are necessary to achieve a biological response and, as indicated above in the section on antagonists, the inability of the steroid-receptor complex to act anywhere in the sequence will prevent expression of the hormonal effect.

CONCLUDING REMARKS

Current understanding of receptor mechanisms provides insight into several aspects of adrenal steroid action. Explanations have been offered to rationalize the following phenomena: (*a*) steroid specificity, (*b*) overlapping physiological actions of mineralocorticoids and glucocorticoids, (*c*) differences in potency of various classes of steroids, (*d*) the mechanism of action of antagonists, and (*e*) pathways for the development of steroid resistance at a cellular level. Further understanding of receptors may clarify the role of steroids in differentiation, shed light on additional functions of steroids at tissue sites, and enable the design of new and more potent steroid antagonists. Ultimately, perhaps, continued exploration of the steroid-receptor interaction may elucidate hormonal mechanisms of gene regulation in animal cells.

Literature Cited

1. Raspe, G, Ed. 1970. Schering workshop on steroid hormone receptors. *Advances in the Biosciences 7.* Oxford: Pergamon
2. Litwack, G. 1970. *Biochemical Actions of Hormones.* New York: Academic
3. O'Malley, B. W. 1971. Mechanisms of action of steroid hormones. *New Engl. J. Med.* 284:370–77.
4. Feldman, D., Funder, J. W., Edelman, I. S. 1972. Subcellular mechanisms in the action of adrenal steroids. *Am. J. Med.* 53:545–60.
5. King, R. J. B., Mainwaring, W. I. P. 1974. *Steroid Receptor Interactions.* London: Butterworth
6. Milgrom, E., Atger, M., Baulieu, E. E. 1973. Studies on estrogen entry into uterine cells and on estradiol-receptor complex attachment to the nucleus—is the entry of estrogen into uterine cells a protein-mediated process? *Biochim. Biophys. Acta* 320:267–83.
7. Ishii, D. M., Pratt, W. B., Aronow, L. 1972. Steady-state level of the specific glucocorticoid binding component in mouse fibroblasts. *Biochemistry* 11: 3896–3904
8. Munck, A. et al 1972. Glucocorticoid-receptor complexes and the earliest steps in the action of glucocorticoids on thymus cells. *J. Steroid Biochem.* 3:567–78
9. Marver, D., Goodman, D., Edelman, I. S. 1972. Relationships between renal cytoplasmic and nuclear aldosterone receptors. *Kidney Int.* 1:210–23
10. Rousseau, G. G., Baxter, J. D., Tomkins, G. M. 1972. Glucocorticoid receptors: Relations between steroid binding and biological effects. *J. Mol. Biol.* 67:99–142.
11. Rousseau, G. G., Baxter, J. D., Higgins, S. J., Tomkins, G. M. 1973. Steroid-induced nuclear binding of glucocorticoid receptors in intact hepatoma cells. *J. Mol. Biol.* 79:539–54
12. Marver, D., Stewart, J., Funder, J. W., Feldman, D., Edelman, I. S. 1974. Renal aldosterone receptors: Studies with [^{3}H] aldosterone and the anti-mineralocorticoid [^{3}H] spirolactone (SC-26304). *Proc. Nat. Acad. Sci. USA* 71:1431–35
13. Turnell, R. W., Kaiser, N., Milholland, R. J., Rosen, F. 1974. Glucocorticoid receptors in rat thymocytes. Interactions with the antiglucocorticoid cortexolone and mechanism of its action. *J. Biol. Chem.* 249:1133–38
14. Wira, C., Munck, A. 1974. Glucocorticoid-receptor complexes in rat thymus cells: "Cytoplasmic"-nuclear transformations. *J. Biol. Chem.* 249:5328–36
15. Swaneck, G. E., Highland, E., Edelman, I. S. 1969. Sterospecific nuclear and cytosol aldosterone-binding proteins of various tissues. *Nephron* 6:297–316
16. Funder, J. W., Feldman, D., Edelman, I. S. 1972. Specific aldosterone binding in rat kidney and parotid. *J. Steroid Biochem.* 3:209–18
17. Baxter, J. D. et al 1972. Role of DNA and specific cytoplasmic receptors in glucocorticoid action. *Proc. Nat. Acad. Sci. USA* 69:1892–96
18. Higgins, S. J., Rousseau, G. G., Baxter, J. D., Tomkins, G. M. 1973. Nature of nuclear acceptor sites for glucocorticoid and estrogen-receptor complexes. *J. Biol. Chem.* 248:5873–79
19. Munck, A., Brinck-Johnsen, T., 1974. Is cortisol metabolized as it dissociates from glucocorticoid receptors in thymus cells? *J. Steroid Biochem.* 5:203–5
20. Funder, J. W., Feldman, D., Edelman, I. S. 1973. The roles of plasma binding and receptor specificity in the mineralocorticoid action of aldosterone. *Endocrinology* 92:994–1004
21. Rousseau, G., Baxter, J. D., Funder, J. W., Edelman, I. S., Tomkins, G. M. 1972. Glucocorticoid and mineralocorticoid receptors for aldosterone. *J. Steroid Biochem.* 3:219–27
22. Funder, J. W., Feldman, D., Edelman, I. S. 1973. Glucocorticoid receptors in the rat kidney: The binding of tritiated-dexamethasone. *Endocrinology* 92: 1005–13
23. Feldman, D., Funder, J. W., Edelman, I. S. 1973. Evidence for a new class of corticosterone receptors in rat kidney. *Endocrinology* 92:1429–41
24. Duval, D., Funder, J. W. 1974. The binding of tritiated aldosterone in the rat liver cytosol. *Endocrinology* 94: 575–79
25. Munck, A., Wira, C. 1971. Glucocorticoid receptors in rat thymus cells. *Advan. Biosc.* 7:301–30
26 Hackney, J. F., Gross, S. R., Aronow, L., Pratt, W. B. 1970. Specific glucocorticoid-binding macromolecules from mouse fibroblasts growing in vitro. A possible steroid receptor for growth inhibition. *Mol. Pharmacol.* 6:500–12

27. Beato, M., Feigelson, P. 1972. Glucocorticoid-binding proteins of rat liver cytosol. I. Separation and identification of the binding proteins. *J. Biol. Chem.* 247:7890–96
28. Baxter, J. D., Tomkins, G. M. 1971. Specific cytoplasmic glucocorticoid hormone receptors in hepatoma tissue culture cells. *Proc. Nat. Acad. Sci. USA* 68:932–37
29. McEwen, B. S., Magnus, C., Wallach, G. 1972. Soluble corticosterone—binding macromolecules extracted from rat brain. *Endocrinology* 90:217–26
30. Watanabe, H., Orth, D. N., Toft, D. D. 1973. Glucocorticoid receptors in pituitary tumor cells. I. Cytosol receptors. *J. Biol. Chem.* 248–7625–30
31. Chader, G. J., Reif-Lehrer, L. 1972. Hormonal effects on the neural retina: Corticoid uptake, specific binding and structural requirements for the induction of glutamine synthetase. *Biochim. Biophys. Acta* 264:186–96
32. Samuels, H. H., Tomkins, G. M. 1970. Relation of steroid structure to enzyme induction in hepatoma tissue culture cells. *J. Mol. Biol.* 52:57–74
33. Feldman, D., Dziak, R., Koehler, R., Stern, P. 1975. Cytoplasmic glucocorticoid binding proteins in bone cells. *Endocrinology* 95:29–36
34. Funder, J. W., Duval, D., Meyer, P. 1973. Cardiac glucocorticoid receptors: The binding of tritiated dexamethasone in rat and dog heart. *Endocrinology* 93:1300–08
35. Giannopoulos, G. 1973. A specific glucocorticoid binding macromolecule of rabbit uterine cytosol. *Biochem. Biophys. Res. Commun.* 54:600–6
36. Ballard, P. L., Baxter, J. D., Higgins, S. J., Rousseau, G. G., Tomkins, G. M. 1974. General presence of glucocorticoid receptors in mammalian tissues. *Endocrinology* 94:998–1002
37. Ballard, P. L., Ballard, R. A. 1972. Glucocorticoid receptors and the role of glucocorticoids in fetal lung development. *Proc. Nat. Acad. Sci. USA* 69:2668–72
38. Giannopoulos, G. 1973. Glucocorticoid receptors in lung. I. Specific binding of glucocorticoids to cytoplasmic components of rabbit fetal lung. *J. Biol. Chem.* 248:3876–83
39. Giannopoulos, G. 1974. Variations in the levels of cytoplasmic glucocorticoid receptors in lungs of various species at different developmental stages. *Endocrinology* 94:450–58
40. Ballard, P. L., Ballard, R. A. 1974. Cytoplasmic receptor for glucocorticoids in lung of the human fetus and neonate. *J. Clin. Invest.* 53:477–86
41. Feldman, D. 1974. Ontogeny of rat hepatic glucocorticoid receptors. *Endocrinology* 95:1199–1207
42. Mulay, S., Giannopoulos, G., Solomon, S. 1973. Corticosteroid levels in the mother and fetus of the rabbit during gestation. *Endocrinology* 93:1342–48
43. Funder, J. W., Feldman, D., Highland, E., Edelman, I. S. 1974. Molecular modifications of anti-aldosterone compounds: Effects on affinity for renal aldosterone receptors. *Biochem. Pharmacol.* 23:1493–1501
44. Kirkpatrick, A. F., Milholland, R. J., Rosen, F. 1971. Stereospecific glucocorticoid binding to subcellular fractions of the sensitive and resistant lymphosarcoma P1798. *Nature New Biol.* 232:216–18
45. Rosenau, W., Baxter, J. D., Rousseau, G. G., Tomkins, G. M. 1972. Mechanism of resistance to steroids: Glucocorticoid receptor defect in lymphoma cells. *Nature New Biol.* 237:20–24
46. Lippman, M. E., Halterman, R. H., Leventhal, B. G., Perry, S., Thompson, E. B. 1973. Glucocorticoid-binding proteins in human acute lymphoblastic leukemic blast cells. *J. Clin. Invest.* 52:1715–25
47. Gailani, S. et al 1973. Specific glucocorticoid binding in human hemopoietic cell lines and neoplastic tissue. *Cancer Res.* 33:2653–57

Copyright 1975. All rights reserved

REGULATION OF ALBUMIN METABOLISM[1]

❖7134

Marcus A. Rothschild, M.D., Murray Oratz, Ph.D.,
and Sidney S. Schreiber, M.D.
The Radioisotope Service, New York Veterans Administration Hospital, New York, NY 10010; and the Department of Medicine, New York University School of Medicine, and the Department of Biochemistry, New York University College of Dentistry, New York, NY 10016

INTRODUCTION

Serum albumin has a long history spanning millions of years, during which time changes in the molecule have paralleled man's evolution from his ancestors (1, 2). Albumin or albumin-like proteins have been observed in nearly all species though cross-reactivity between species and with man is not noted (1, 3, 4). The albumin molecule in man has been observed as early as the first month of gestation (5, 6). It may not exist as a unique single chain polymer within the plasma since specific variants have been reported where the probable difference is due to a change in a single amino acid in the 575 that constitute the albumin molecule (1, 3). It is recognized that most of the biochemical differences or electrophoretic behavior of the albumin depend upon the substances that are bound to albumin (7). Albumin is a carrier for practically all metabolites, ions, amino acids, drugs, hormones, and products of toxic degradation (1, 8, 9). As the number of these bound moieties, such as fatty acids, changes, conformational or charge changes may occur and electrophoretic mobility alters as a result. Albumin also serves as the prime regulator of osmotic pressure (10). It is fascinating that a molecule with so many diverse properties, and one which normally makes up the largest fraction of the serum proteins, does not seem to be vital, at least in the few patients studied who are born without any albumin (11–13). It is possible, as with so many other clinical conditions, that when hypoalbuminemia develops rapidly due to loss or disease, the body finds it difficult to adjust to this new situation. But, when confronted with a chronic disorder, even the absence of a relatively vital substance can be tolerated with ease.

---

[1] Supported in part by the National Institutes of Health Grant AA 00959 and the Bear Fund.

ALBUMIN DISTRIBUTION

Albumin is widely distributed not only in plasma and interstitial fluid, but in all bodily secretions such as feces, bile, tears, sweat, bronchial secretion, exudates, and transudates (8, 14). Most available data indicate that approximately 40% of the exchangeable albumin pool is normally located within the plasma space and 60% in extraplasma tissues (1, 8, 15–17). The skin and muscle serve as the largest storehouse for extravascular albumin with only small quantities of albumin located within the internal viscera and liver. The major bulk of extravascular albumin is in the interstitial spaces of the tissues and possibly in the smaller lymphatic radicals. The larger lymphatic ducts contain less than 2–3% of the extravascular albumin. Control of this complex distribution scheme must be separate from the regulation of the rapid movement of water and electrolytes which takes place between various body compartments. The interstitial space is made up of a complex polysaccharide protein matrix through which the large protein molecules move with difficulty (18, 19). In fact, molecules such as albumin find available for their solution only a small fraction of the interstitial space which is available to smaller molecules for solubility. Albumin enters this interstitial space probably through endothelial vesicles which permit the passage of molecules of only certain molecular size, whereas water and other solutes exchange quickly. Normally, over one half of the extraplasma albumin spends at least two days in this pool and as much as one fifth requires at least ten days for transit. Fibrinogen, for example, has only 10–15% of its total exchangeable pool located in extravascular sites, while α-2-macroglobulin with a molecular weight of greater than 800,000 is nearly completely excluded from interstitial fluid. Thus, the entrance of protein into the interstitial space may well be controlled by the size of the vesicles and the permeability of the basement membrane (15). Other tissues which come into equilibrium faster than do muscle or skin, such as the liver, have an insignificant quantity of interstitial albumin. The liver, for example, contains less than 1% of the total exchangeable albumin even though it is the site of albumin synthesis (20, 21). Hepatic interstitial albumin is also limited to a fraction of the available interstitial water. Rapid turnover of hepatic albumin would not be detrimental to the stability of the system. However, rapid turnover of the large mass of interstitial albumin in the extravascular spaces of the skin and muscle would pose a significant problem to the maintenance of the integrity of the intravascular pool, and, of course, blood volume. In stress situations, e.g. burns, trauma, and infection, the extravascular albumin rapidly becomes available to augment the plasma pool, indicating a change in this normally stable extra plasma pool. It is important to recognize that the extravascular pool is now in rapid equilibrium with the plasma. Perhaps changes within the interstitial matrix of skin and muscle occur first, permitting the rapid outflow from this albumin pool to supplement the plasma. Ascitic albumin, which may account for more exchangeable albumin than in the rest of the body, equilibrates very slowly—a month or more is not unusual (22). A direct loss of newly synthesized albumin from the liver into ascites bypassing the systemic circulation and the thoracic duct lymph has been demonstrated (23).

ALBUMIN SYNTHESIS

Mechanism

Within the liver there are a series of complex intracellular organelles responsible for protein synthesis (24–32), and only an overview of this rapidly changing field will be presented here. Serum albumin is felt to be synthesized primarily on a polysome consisting of mRNA and rRNA bound to the microtubular structure, the endoplasmic reticulum (25, 26), although in isolated systems some free polysomes are capable of binding to an anti-albumin antibody (33, 34). Futher, using a reticulocyte system, it was possible to show that exogenous mRNA for albumin could be translated by free polysomes (35).

Cellular Transport

The scenario for the liver's synthesis of albumin is similar to that for other secreted proteins. Following initiation, activated amino acids are brought to the site of the connection between the smaller ribosomal subunit and the mRNA; the growing peptide is translocated from one spot to another and the albumin molecule inserted through the larger of the ribosomal subunits (60S) into the endoplasmic reticulum. Synthesis is terminated and the completed albumin molecule is transported through the endoplasmic reticulum to the Golgi apparatus to be extruded in some fashion into hepatic plasma. This process requires 15–20 min for completion and the integrity of the endoplasmic reticulum through which albumin must pass may well be one of the factors regulating the rate of albumin production (24–32, 36, 37). The second system within the liver, the free polysome system, is not normally responsible for albumin production in vivo. This system, unattached to the endoplasmic reticulum, mainly synthesizes protein for cellular consumption in vivo. This difference is important because the bound and free polysomes respond in separate ways to various stresses applied to the cell.

Newly synthesized albumin bypasses the lymph and hepatic interstitial fluid and is essentially completely dumped into hepatic plasma (38, 39). That fraction which reaches hepatic lymph more or less directly is equivalent to the relative rates of hepatic plasma and lymph flow and is approximately 0.05%. In isolated perfused systems, the total albumin radioactivity is released into this closed circuit (from a rapid label) and requires some 60–90 min, much longer than the 20 min or so for the synthesis and transport of the initial albumin molecule. It is conceivable, as has been suggested, that there is a precursor albumin molecule (40) synthesized by the liver which remains dormant until release is called for; alternatively, as with other systems, there may be a host of transit pathways, some requiring longer transit times than other. These questions have not as yet been settled. Cellular potassium levels play a role in albumin transport in liver slice studies, although usually this preparation is difficult to equate to one where the circulation is intact (41). Assuming that the 20-min delay for albumin release holds true for a full 24 hr and that at any one moment there is 0.4 mg (T. Peters Jr, personal communication) of albumin per gram of liver awaiting release, then in the 1500-g liver of an "average man" of 70 kg, albumin synthesis would be about 43 g per day, or 3 times the normal rate. Thus most of the liver cells must not be making (or capable of making) albumin at any one time, or alternatively, albumin synthesis is capable of rapid changes in rate. Evidence is available that one third of the liver cells stain for albumin at any one moment (42), and further that the hepatic protein synthesizing system is very sensitive to aggregation and disaggregation secondary to transient fluxes of gut-delivered amino acids (43–49). Albumin synthesis normally varies between 150–200 mg/kg/day (Table 1). Albumin synthesis is reported to be enhanced shortly after birth and to remain high in the early neonatal period (50, 51). Further, albumin synthesis may be increased in older rats compared to younger rats (although it may not be as responsive to stress) (52). Control over the excess production of albumin must be very strict for normal or diseased individuals do not show hyperalbuminemia.

FACTORS INFLUENCING ALBUMIN SYNTHESIS

Nutrition

The factors which interdigitate to effect control over albumin synthesis in vivo include: nutrition, hormonal milieu, environment, protein concentration pool size and distribution, and disease. The most important factor regulating albumin production is nutrition (43–49, 53–56). Adequate nitrogen intake is probably basic to all mechanisms involved in the regulation of protein metabolism—not only of serum albumin. Adaptation to altered nitrogen intake is one of the most complex yet vital reactions of the whole body, and serum albumin synthesis is nearly instantaneously decreased when exogenous administration of amino acids is lowered, due either to a decreased intake, intestinal destruction of amino acids, altered absorption, or changes in blood flow to the liver. During the first few days on a protein-deficient diet or starvation, RNA rapidly disappears from the liver cell due initially to an increase in RNA degradation associated with a decrease in RNA synthesis. It has been suggested that amino acid supply regulates polysome aggregation, and as monosomes increase, RNase increases causing accelerated RNA breakdown (45). As the fast or protein depletion continues, the monosomal fraction no longer accumulates and the total RNA stabilizes at a new level without further excessive loss of RNA. This is a rapid cycle and probably occurs frequently during the day. The

Table 1 Normal albumin metabolism[a]

Subject of Study	Serum Albumin Level (g/100 ml)	Albumin Synthesis (mg/kg/day)	Albumin Degradation (% plasma albumin pool/day)	Exchangeable Albumin Pool (g/kg)	Intravascular Albumin (%)
Adults:					
Male	3.5–4.5	120–200	6–10	4.0–5.0	38–45
Female	3.5–4.5	120–150	6–10	3.5–4.5	38–45
Children:					
13 days–14 mo	3.3–4.3	180–300	10–11	6.0–8.0	33–43
3–8 yr	4.2–5.0	130–170	6–9	3.0–4.0	46–51

[a] Reprinted with permission from the *New England Journal of Medicine* (1972, Vol. 286).

amino acid level is the most sensitive regulator of RNA metabolism, for ATP and GTP levels have to decrease by 60% or so to effect polysomal disaggregation per se (57–59). The large loss of ribosomal RNA from the cell must lower the available ribosomal subunits, and it is of interest that the remaining ribosomes are primarily directed towards the free polysome, whereas the endoplasmic-membrane-bound polysome responsible for protein synthesis for export is disaggregated (60). Upon re-administration of an adequate diet, or of specific amino acids, particularly tryptophan, RNA synthesis increases, the bound polysome re-aggregates, and albumin production returns to values that may be even higher than those found in the fed state (60–64). This situation is probably the result of the fact that the mRNA for albumin is rather long lived or stable, and excess albumin might be expected to be produced until refeeding results in a newer supply of shorter lived mRNAs to compete with the available ribosomal pool (44). It is interesting to realize that in situations of starvation or malnutrition as much as 90% of the amino acids available for the synthesis of protein comes from recycling intrahepatocellular amino acids and not from extrahepatic sources (65). The lack of charging of specific tRNAs may well be the specific event leading to disaggregation, which in turn precedes RNA loss (66). Why the endoplasmic-membrane-bound polysome should be so sensitive to the external source of the amino acids is not clear.

Hormones

The action of hormones, particularly thyroid and cortisone, appear to be primarily directed towards RNA metabolism (67–69). Lack of thyroid hormone results in lowered mRNA and rRNA synthesis, and the addition of thyroid hormone reverses these findings with an increase in nRNA, followed by rRNA polymerase and later by mRNA polymerase. Recent studies have indicated that an early effect of thyroid hormone is to stimulate the rate of peptide chain elongation (70). In vivo excess thyroid hormone results in enhanced albumin synthesis (71, 72).

Cortisone also affects RNA synthesis, with eventual stimulation of tRNA, mRNA, nRNA, and unclassified RNAs (73–77). Nitrogen accumulates in the liver as long as the diet is adequate (78, 79). Cortisone may promote binding of ribosomes to the endoplasmic reticulum. The synthesis of nRNA may even be slowed initially and achieve higher values later on (80), which may help explain the observation that albumin synthesis may be slowed initially after cortisol (81). However, in vivo and in vitro studies agree that cortisone eventually results in a stimulation of albumin synthesis (82–85).

Other hormones also affect hepatic RNA synthesis. Growth hormone promotes the synthesis of RNA and stimulates protein synthesis in vitro (86–88). Albumin synthesis has been stimulated in vivo by growth hormone (89, 90). The lack of insulin in experimental animals results in a loss of endoplasmic reticulum (91) and protein synthesis has been shown to be stimulated by insulin (84, 92). Clinical diabetes is not necessarily associated with a lowered albumin synthetic rate. However, in all likelihood, a "normal" balance of hormones is probably the best milieu for effective protein synthesis and this effect is probably mediated again via an adequate supply of amino acids.

Osmotic Equilibrium

Another important factor in terms of the regulation of albumin synthesis is oncotic pressure (93–98). Reciprocal depressions of the serum albumin level in the face of elevations in the globulin content of the plasma have long been observed, and it has been clearly shown in vivo and in vitro that as the colloidal content is raised either by dextran, gamma globulin, or albumin (the latter only in the isolated perfused system) (95), the absolute rates of albumin synthesis decrease. When exogenous albumin is infused in vivo, the accumulation of excess colloid is prevented not by changes in synthesis but by marked increases in albumin degradation. These observations indicate that albumin synthesis and degradation are not interdependent, and, furthermore, cast doubt on the efficacy of large infusions of salt-poor human serum albumin in significantly altering the exchangeable albumin pool. Most of the intravenously administered albumin is either lost rapidly in the urine or degraded. Changes in albumin synthesis are related inversely to extravascular albumin and particularly within the interstitial space of the liver. Whereas the interstitial volume of the hepatic tissue amounts to 10–12% of wet liver weight, only one tenth of this space is essentially available for solubility for human serum albumin. This limited domain for albumin space would make a very sensitive sensor for the colloidal regulation of albumin production. Osmotic regulation of albumin production does not seem to be limited to the mammal, for the shark maintains its osmotic equilibrium with the external environment by maintaining a high urea content (98, 99). As the shark is moved to a lower salt content, the level of urea drops, and an albuminoid-like molecule has been noted electrophoretically.

Environment

Environment also plays a role in albumin production, for during experimental acclimatization to heat, albumin synthesis was depressed to less than one third of the normal rate (100). Cold acclimatization failed to alter albumin production while acclimatization to altitude has been reported to stimulate albumin degradation (101, 102). Clinically, in tropical climates there is a tendency for the albumin level to be depressed and the globulin fraction of the blood elevated, though chronic infection and altered nutrition may play a role in these changes.

DISEASE CONSIDERATIONS

Lowered serum albumin levels are classically seen in many disease entities and a host of diseases and syndromes may be associated with decreased albumin production (Table 2). The important factor to remember is that in these disorders probably all of the factors which influence albumin production in the normal state may be markedly altered, and thus the causative relationship between any one factor and a changed albumin production would be hard to evaluate.

Cirrhosis

In cirrhosis of the liver it had been noted that a low serum albumin level was a bad prognostic sign, but this occurred only when the disease was not further complicated

Table 2 Syndromes with frequently decreased albumin production[a]

- Malnutrition
  - Starvation, amino acid deficiencies
  - Protein - calorie imbalance (kwashiorkor)
- Malabsorption
- Cirrhosis
- Stress
  - Surgical
  - Injury and burns
  - Infection
- Carcinoma
- Hypergammaglobulinemia
- Hypothyroidism
- Environment (Acclimatization to heat)
- Analbuminemia
- Alcohol, Carbon tetrachloride, and other hepatic toxins

[a] Reprinted with permission from the *New England Journal of Medicine* (1972, Vol. 286).

by the development of ascites (103). This paradox might be explained by the fact that many of the patients with ascites have elevated exchangeable albumin pools (104–107), and if this is the case, albumin synthesis may not necessarily be impaired. Support for this concept has recently come from studies in cirrhosis with ascites (108) and in cirrhotic subjects under the influence of steroids where albumin metabolism was not found to be necessarily reduced (109, 110). Thus, once the insults of alcohol and poor nutrition are removed, the liver seems capable of normal albumin synthesis.

Gastrointestinal Disease

In gastrointestinal disease there may be loss of plasma proteins through the intestinal tract secondary to altered mucosal membrane, or due to excessive lymphatic loss (111–114). Pathways for albumin loss into intercellular spaces have been demonstrated (115). Albumin synthesis may be low, normal, or elevated (113, 116), and the factor limiting the rate of albumin synthesis is not related to protein loss but to hepatic supply of amino acids.

Renal Disease

Albumin metabolism may be altered in renal disease (117–122). Normally the renal tubules degrade the small amount of albumin delivered to them and the amino acids are recycled. In renal disease, depending on the predominant lesion, excess degradation could take place if the tubules were intact and if large loss of albumin occurred at the glomerulus. Excessive proteinuria could occur at the expense of degradation or a combination of these changes could balance one another. A decrease in functioning renal mass would lead to slower degradation. Albuminuria and thus albumin synthesis is dependent on nitrogen intake. It is only when the diet is adequate and albuminuria marked that albumin synthesis is increased. Whether due to low synthesis or excessive loss, hypoalbuminemia persists in nephrosis. Why the excessive synthetic capacity of the liver for albumin found in vitro (123) in nephrotic animals is not seen in vivo is unclear.

Alcohol and CCl_4

Albumin synthesis can be impeded in the face of toxic substances or metabolites. CCl_4 rapidly disrupts the hepatic endoplasmic reticulum and albumin synthesis is lowered. Alcohol likewise alters the integrity of the endoplasmic reticulum and disrupts the endoplasmic-membrane-bound polysome with a rapid loss in the albumin synthesizing potential (124–129). In isolated perfused liver systems, there is no impediment to the rate of release of preformed albumin (129). Following acute exposure to alcohol, there is only a small loss of hepatic RNA, a disaggregation of the bound polysome, and a loss of albumin synthesizing capacity. The free polysome is minimally disaggregated. Since there is no marked loss of RNA, as seen in fasting where 30% or so may be lost in 24 hr, the disaggregation of the bound polysome cannot be ascribed to a deficiency in ribosomes. Recent studies with ^{3}H-uridine suggest that the effect of alcohol is to impede protein initiation on the bound polysome (60). High levels of amino acids can reverse the acute effects of alcohol in vivo. The hepatic cell returned towards normal as seen on electron microscopy and protein synthesis improved (130). Similar biochemical improvement has been found with specific amino acids using the isolated rabbit liver perfused with alcohol and high levels of amino acids, as long as the liver was derived from a fed donor (60). The combination of stresses, namely fasting and acute exposure to alcohol, prevents the liver from recovery from either stress by procedures which were capable of restimulating albumin production when either stress was applied alone. These observations again raise the question of the relative importance of chronic alcohol exposure and of altered nutrition in predisposing the liver to severe disease.

ALBUMIN DEGRADATION

While much is known about albumin synthesis, and some about albumin distribution, very little is known about albumin degradation. Serum albumin is not a suicidal protein, and in the process of performing transport or oncotic functions does not appear to undergo any significant transformation which would make it susceptible to recognition for degradation as do the sialic acid residues in the metabolism of the glycoproteins (131, 132). Isotopic studies have indicated that the location for the albumin degradative compartment is in close approximation to both the plasma and the extravascular space, but is really not part of either (133). No one organ has been indicted as being particularly responsible for albumin degradation (117, 118, 121, 133–140). A small fraction of the total amount of albumin degraded per day is lost into the stool, excreted and degraded by the kidney, and metabolized by the liver. Some is present, as we have already noted above, in body fluids such as bile, and in other fluids that may be lost from the organism (9, 141). In renal disease, an excess quantity of protein may be lost in the urine or degraded by the tubular mass, and a similar excessive loss may occur through the gastrointestinal tract. Obviously in skin disease, massive burns, or weeping wounds, serum albumin may be lost or subject to degradation in areas of exudate. But these are abnormal situations, and the mechanism by which serum albumin is degraded under normal control conditions is still a complete mystery. The tools for the study of this problem are nowhere as well developed as they are for the study of distribution or synthesis. Labeled proteins are valid tracers for endogenous albumin metabolism, but since the normal turnover rate is only 3–4% of the exchangeable albumin pool, or some 8 or 9% of the plasma pool per day, any single organ studied for a period of a few hours would be expected to contribute only a small fraction of this percentage and thus we are looking for a small number. Until we have better ideas, the degradation of albumin will remain an intriguing mystery.

In 1839, H. Ancell (142) wrote "Albumin is doubtless one of the most important of the animal proximate principles. It is found not only in the serum of the blood but in the lymph, chyle, in the exhilations from surfaces, in the fluid of cellular tissue, in the aqueous and vitrous humors of the eye, in many other animal fluids". Further he stated that "albumin acts sometimes as an acid and a base, and its power of maintaining its combination is so great." He also noted that albumin proportions are "pretty nearly the same in all higher animals, the levels are nearly the same in the sexes and between the ages of 20–60." Thus, serum albumin metabolism has long interested investigators because of its prominent position among the serum proteins. Further studies will most certainly define the regulation of albumin metabolism. The prominent role of nutrition as the basic factor underlying all others is the most likely candidate at present.

Literature Cited

1. Peters, T. Jr. 1970. Serum albumin. *Advan. Clin. Chem.* 13:37–111
2. Sarich, V. M., Wilson, A. C. 1967. Rates of albumin evolution in primates. *Proc. Nat. Acad. Sci. USA* 58:142–48
3. Melartin, L. 1967. Albumin polymorphism in man: studies on albumin variants in North American native populations. *Acta Pathol. Microbiol. Scan. Suppl.* 191:1–50
4. Engle, R. L. Jr., Wood, K. R. 1960. *Comparative Biochemistry and Embroyology in the Plasma Proteins,* ed. F. W. Putnam, 2:183. New York: Academic
5. Gitlin, D., Biasucci, A. 1969. Development of γG, γA γM, β_{ic}/β_{1A}, C'1 Esterase inhibitor, ceruloplasmin, transferrin, hemopexin, haptoglobin, fibrinogen, plasminogen, α_1-antirypsin, orosomucoid, β-lipoprotein, α_2-macroglobulin, and prealbumin in the human conceptus. *J. Clin. Invest.* 48:1433–46
6. Miller, S. A. 1969. Protein metabolism during growth and development. In *Mammalian Protein Metabolism,* ed. H. N. Munro, 3:183–233. New York: Academic. 571 pp.
7. Luetscher, J. A. 1939. Serum albumin II. Identification of more than one albumin in horse and human serum by electrophoretic mobility in acid solution. *J. Am. Chem. Soc.* 61:2888–90
8. Von Bennhold, H. 1962. Uberlick uber die tranport-function der Serumeiweisskorper. In *Protides of the Biological Fluids,* ed. H. Peeters. Amsterdam: Pergamon. 9th ed.
9. Rothschild, M. A., Oratz, M., Schreiber, S. S. 1969. Serum albumin. *Am. J. Dig. Dis.* 14:711–14
10. Schultze, H. E., Heremans, J. F., Eds. 1966 *Molecular Biology of Human Proteins,* 1:600. New York: Elsevier. 903 pp.
11. Bennhold, H., Kallee, E. 1959. Comparative studies on the half-life of I^{131}-labeled albumins and nonradioactive human serum albumin in a case of analbuminemia. *J. Clin. Invest.* 38:863–72
12. Freeman, T. 1969. Analbuminemia—a study of albumin and transferrin metabolism. In *Physiology and Pathophysiology of Plasma Protein Metabolism,* ed.

G. Birke, R. Norberg, L. O. Plantin, 75–85. New York: Pergamon

13. Waldmann, T. A., Gordon, R. S. Jr., Rosse, W. 1964. Studies on the metabolism of the serum proteins and lipids in a patient with analbuminemia. *Am. J. Med.* 37:960–68
14. See Ref. 10, pp. 783, 791, 806, 821, 829, 853, 865, 874, 877
15. Reeve, E. B., Chen, A. Y. 1970. Protein distribution regulation of interstitial albumin. In *Plasma Protein Metabolism,* ed. M. A. Rothschild, T. Waldmann, 89–109. New York: Academic. 560 pp.
16. Rosenoer, V. M., Rothschild, M. A. 1970. The extravascular transport of albumin. See Ref. 15, pp. 111–27
17. Katz, J., Sellers, A. L., Bonorris, G., Golden, S. 1970. Studies on the extravascular ablumin of rats. See Ref. 15, pp. 129–54
18. Laurent, T. C., Ogston, A. G. 1963. The interaction between polysaccharides and other macromolecules. IV. The osmotic pressure of mixtures of serum albumin and hyaluronic acid. *Biochem. J.* 89:249–53
19. Ogston, A. G., Phelps, C. F. 1961. The partition of solutes between buffer solutions and solutions containing hyaluronic acid. *Biochem. J.* 78:827–33
20. Dewey, W. C. 1958. Amount of plasma proteins and red cells and the vascular extravascular exchange of plasma proteins in the rat as measured with I^{131} and Cr^{51} labels. *Univ. Rochester Atomic Energy Rep. No. 524*
21. Rothschild, M. A., Oratz, M., Evans, C. E., Schreiber, S. S. 1966. Role of hepatic interstitial albumin in regulating albumin synthesis. *Am. J. Physiol.* 210:57–62
22. Berson, S. A., Yalow, R. S. 1954. The distribution of I^{131}-labeled human serum albumin introduced into ascitic fluid: analysis of the kinetics of a three compartment catenary transfer system in man and speculations on possible sites of degradation. *J. Clin. Invest.* 33:377–87
23. Zimmon, D. S., Oratz, M., Kessler, R., Schreiber, S. S., Rothschild, M. A. 1969. Albumin to ascites: Demonstration of a direct pathway bypassing the systemic circulation. *J. Clin. Invest.* 48:2074–78
24. Haselkorn, R., Rothman-Denes, L. B. 1973. Protein synthesis. *Ann. Rev. Biochem.* 42:397–438
25. Henshaw, E. C., Hirsch, C. A., Morton, B. E., Hiatt, H. H. 1971. Control of protein synthesis in mammalian tissues through changes in ribosome activity. *J. Biol. Chem.* 246:436–46
26. Munro, H. N. 1970. A general survey of mechanisms regulating protein metabolism in mammals. In *Mammalian Protein Metabolism,* ed. H. N. Munro, 4:3–130. New York: Academic
27. Redman, C. M. 1969. Biosynthesis of serum proteins and ferritin by free and attached ribosomes of rat liver. *J. Biol. Chem.* 244:4308–15
28. Lengyel, P., Soll, D. 1969. Mechanism of protein biosynthesis. *Bacteriol. Rev.* 33:264–301
29. Munro, A. J., Jackson, R. J., Korner, A. 1964. Studies on the nature of polysomes. *Biochem. J.* 92:289–99
30. Blobel, G., Sabatini, D. D. 1970. Controlled proteolysis of nascent polypeptides in rat liver cell fractions. 1. Location of the polypeptides within ribosomes. *J. Cell Biol.* 45:130–57
31. Sabatini, D. C., Tashiro, Y., Palade, G. E. 1966. On the attachment of ribosomes to microsomal membranes. *J. Mol. Biol.* 19:503–24
32. Falvey, A. K., Staehlein, T. 1970. Structure and function of mammalian ribosomes. I. Isolation and characterization of active liver ribosomal subunits. *J. Mol. Biol.* 53:1–19 Ibid. Exchange of ribosomal subunits at various stages of in vitro polypeptide synthesis. 53:21–34
33. Ikehara, Y., Pitot, H. C. 1973. Localization of polysome-bound albumin and serine dehydratase in rat liver cell fractions. *J. Cell Biol.* 59:28–44
34. Uenoyama, K., Ono, T. 1972. Synthesis of albumin by the free polyribosomes in hepatoma. *Biochim. Biophys. Acta* 281:124–29
35. Shafritz, D. A. 1973. Evidence for nontranslated messenger ribonucleic acid in membrane-bound and free polysomes of rabbit liver. *J. Biol. Chem.* 249:89–93
36. Peters, T. Jr., Fleisher, B., Fleischer, S. 1971. The biosynthesis of rat serum albumin. IV. Apparent passage of albumin through the golgi apparatus during secretion. *J. Cell Biol.* 246:240–44
37. Glaumann, H., Ericsson, J. L. E. 1970. Evidence for the participation of the Golgi apparatus in the intracellular transport of nascent albumin in the liver cell. *J. Cell Biol.* 47:555–67
38. Woolley, G., Courtice, F. C. 1962. The origin of albumin in hepatic lymph. *Aust. J. Exp. Biol. Med. Sci.* 40:121–28
39. Smallwood, R. A., Jones, E. A., Craigie, A., Raia, S., Rosenoer, V. M. 1968. The

delivery of newly synthesized albumin and fibrinogen to the plasma in dogs. *Clin. Sci.* 35:35–43

40. Judah, J. D., Gamble, M., Steadman, J. H. 1973. Biosynthesis of serum albumin in rat liver. *Biochem. J.* 134:1083–91
41. Judah, J. D., Nicholls, M. R. 1970. Role of liver-cell potassium ions in secretion of serum albumin and lipoproteins. *Biochem. J.* 116:663–69
42. Feldmann, G., Penaud-Laurencin, J., Crassous, J., Benhamou, J. P. 1972. Albumin synthesis by human liver cells: Its morphological demonstration. *Gastroenterology* 63:1036–48
43. Munro, H. N., Waddington, S., Begg, D. J. 1965. Effect of protein intake on ribonucleic acid metabolism in liver cell nuclei of the rat. *J. Nutr.* 85:319–28
44. Wilson, S. H., Hill, H. Z., Hoagland, M. B. 1967. Physiology of rat-liver polysomes:protein synthesis by stable polysomes. *Biochem. J.* 103:567–72
45. Enwonwu, C. O., Munro, H. N. 1970. Rate of RNA turnover in rat liver in relation to intake of protein. *Arch. Biochem. Biophys.* 138:532–39
46. Rothschild, M. A., Oratz, M., Mongelli, J., Schreiber, S. S. 1968. Effects of a short-term fast on albumin synthesis studied in vivo, in the perfused liver, and on amino acid incorporation by hepatic microsomes. *J. Clin. Invest.* 47:2591–99
47. Baliga, B. S., Pronczuk, A. W., Munro, H. N. 1968. Regulation of polysome aggregation in a cell-free system through amino acid supply. *J. Mol. Biol.* 34: 199–218
48. Staehelin, T., Verney, E., Sidransky, H. 1967. The influence of nutritional change on polyribosomes of the liver. *Biochim. Biophys. Acta* 145:105–19
49. Wannemacher, R. W. Jr., Cooper, W. K., Yatvin, M. B. 1968. The regulation of protein synthesis in the liver of rats: mechanisms of dietary amino acid control in the immature animal. *Biochem. J.* 107:615–23
50. Yeoh, G. C. T. 1973. Albumin and transferrin synthesis in the liver of rats in the immediate postnatal period. *Biochim. Biophys. Acta* 331:421–29
51. Wise, R. W., Oliver, I. T. 1967. Plasma albumin synthesis during neonatal development of the rat. *Biochem. J.* 102:760–62
52. Ove, P., Obenrader, M., Lansing, A. 1972. Synthesis and degradation of liver proteins in young and old rats. *Biochim. Biophys. Acta* 277:211–21
53. Waterlow, J. C. 1968. Observations on the mechanism of adaptation to low protein intakes. *Lancet.* 2:1091–97
54. Cohen, S., Hansen, J. D. L. 1962. Metabolism of albumin and γ-globulin in kwashiorkor. *Clin. Sci.* 23:351–59
55. Hoffenberg, R., Black, E., Brock, J. F. 1966. Albumin and γ-globulin tracer studies in protein depletion states. *J. Clin. Invest.* 45:143–52
56. Kirsch, R., Frith, L., Black, E., Hoffenberg, R. 1968. Regulation of albumin synthesis and catabolism by alteration of dietary protein. *Nature (London)* 217:578–79
57. Oler, A., Farber, E., Shull, K. H. 1969. Resistance of liver polysome ATP deficiency in the male rat. *Biochim. Biophys. Acta* 190:161–69
58. Puddu, P., Caldarer, C. M., Marchetti, M. 1967. Studies on ethionine-induced fatty liver. *Biochem. J.* 102:163–67
59. Clifford, B. J., Riumallo, J. A., Baliga, B. S., Munro, H. N., Brown, P. R. 1972. Liver Nucleotide metabolism in relation to amino acid supply. *Biochim. Biophys. Acta* 277:443–58
60. Rothschild, M. A., Oratz, M., Schreiber, S. S. 1974. Alcohol, amino acid and albumin synthesis. *Gastroenterology,* Vol. 67. In press
61. Rothschild, M. A., Oratz, M., Mongelli, J., Schreiber, S. S. 1969. Amino acid regulation of albumin synthesis. *J. Nutr.* 98:395–403
62. Henderson, A. R. 1970. The effect of feeding with a tryptophan-free amino acid mixture on rat liver magnesium ion-activated deoxyribonucleic acid-dependent ribonucleic acid polymerase. *Biochem. J.* 120:205–14
63. Sidransky, H., Sarma, D. S. R., Bongiorno, M., Verney, E. 1968. Effect of dietary tryptophan on hepatic polyribosomes and protein synthesis in fasted mice. *J. Biol. Chem.* 243:1123–32
64. Kelman, L. et al 1972. The effects of amino acids on albumin synthesis by the isolated perfused rat liver. *Biochem. J.* 129:805–9
65. Gan, J. C., Jeffay, H. 1967. Origins and metabolism of the intracellular amino acid pools in rat liver and muscle. *Biochim. Biophys. Acta.* 148:448–59
66. Allen, R. E., Raines, P. L., Regen, D. M. 1969. Regulatory significance of transfer RNA charging levels. I. Measurements of charging levels in livers of chow-fed rats, fasting rats, and rats fed

balanced or imbalanced mixtures of amino acids. *Biochim. Biophys. Acta* 190:323–36

67. Garren, L. D., Richardson, A. P. Jr., Crocco, R. M. 1967. Studies on the role of ribosomes in the regulation of protein synthesis in hypophysectomized rats. *J. Biol. Chem.* 242:650–56
68. Tata, J. R., Widnell, C. C. 1966. Ribonucleic acid synthesis during the early action of thyroid hormones. *Biochem. J.* 98:604–20
69. Widnell, C. C., Tata, J. R. 1966. Additive effects of thyroid hormone, growth hormone and testosterone on deoxyribonucleic acid-dependent ribonucleic acid polymerase in rat-liver nuclei. *Biochem. J.* 98:621–29
70. Mathews, R. W., Oronsky, W. A., Haschemeyer, A. E. V. 1973. Effect of thyroid hormone on polypeptide chain assembly kinetics in liver protein synthesis in vivo. *J. Biol. Chem.* 248:1329–33
71. Rothschild, M. A., Baumann, A., Yalow, R. S., Berson, S. A. 1957. The effect of large doses of desiccated thyroid on the distribution and metabolism of albumin-I^{131} in euthyroid subjects. *J. Clin. Invest.* 36:422–28
72. Lewallen, C. G., Rall, J. E., Berman, M. 1959. Studies of iodoalbumin metabolism. II. The effects of thyroid hormone. *J. Clin. Invest.* 38:88–101
73. Cox, R. F., Mathias, A. P. 1969. Cytoplasmic effects of cortisol in liver. *Biochem. J.* 115:777–87
74. Drews, J., Brawerman, G. 1967. Alterations in the nature of ribonucleic acid synthesized in rat liver during regeneration and after cortisol administration. *J. Biol. Chem.* 242:801–8
75. Yatvin, M. B., Wannemacher, R. W. Jr. 1965. Action of adrenal corticoids on protein metabolism in the thiouracil-treated rat. *Endocrinology* 76:418–26
76. Enwonwu, C. O., Munro, H. N. 1971. Changes in liver polyribosome patterns following administration of hydrocortisone and actinomycin D. *Biochim. Biophys. Acta.* 238:264–76
77. Agarwal, M. K., Hanoune, J., Yu, F. L., Weinstein, I. B., Feigelson, P. 1969. Studies on the effect of cortisone on rat liver transfer ribonucleic acid. *Biochm. J.* 8:4806–13
78. Silber, R. H., Porter, C. C. 1953. Nitrogen balance, liver protein repletion and body composition of cortisone treated rats. *Endocrinology.* 52:518–25
79. Kaplan, S. A., Shimizu, C. S. 1962. Free amino acid and amine concentrations in liver: effects of hydrocortisone and fasting. *Am. J. Physiol.* 202:695–98
80. Schmid, W., Sekeris, C. E. 1972. Sequential stimulation of extranucleolar and nucleolar RNA synthesis in rat liver by cortisol. *FEBS Lett.* 26:109–12
81. Jeejeebhoy, K. N., Bruce Robertson, A., Ho, J., Sodtke, U. 1972. The effect of cortisol on the synthesis of rat plasma albumin fibrinogen and transferrin. *Biochem. J.* 130:533–38
82. Rothschild, M. A. et al 1958. The effects of adrenocortical hormones on albumin metabolism studied with albumin-I^{131}. *J. Clin. Invest.* 37:1229–35
83. Sterling, K. 1960. The effect of Cushing's syndrome upon serum albumin metabolism. *J. Clin. Invest.* 39:1900–8
84. John, D. W., Miller, L. L. 1970. Regulation of net biosynthesis of serum albumin and acute phase proteins: induction of enhanced net synthesis of fibrinogen, α_1-acid glycoprotein, α_2(acute phase)-globulin, and haptoglobin by amino acids and hormones during perfusion of the isolated normal rat liver. *J. Biol. Chem.* 244:6134–42
85. Bancroft, F. C., Levine, L., Tashjian, A. H. Jr. 1969. Serum albumin production by hepatoma cells in culture: direct evidence for stimulation by hydrocortisone. *Biochem. Biophys. Res. Commun.* 37:1028–35
86. Tata, J. R., Williams-Ashman, H. G. 1967. Effects of growth hormone and triiodothyronine on amino acid incorporation by microsomal subfractions from rat liver. *Eur. J. Biochem.* 2:366–74
87. Korner, A. 1964. Regulation of the rate of synthesis of messenger ribonucleic acid by growth hormone. *Biochem. J.* 92:449–56
88. Jefferson, L. S., Korner, A. 1967. A direct effect of growth hormone on the incorporation of precursors into proteins and nucleic acids of perfused rat liver. *Biochem. J.* 104:826–32
89. Ulrich, F., Tarver, H., Li, C. H. 1954. Effects of growth and adrenocorticotropic hormones on the metabolism of albumin in hypophysectomized rats. *J. Biol. Chem.* 209:117–25
90. Kernoff, L. M., Pimstone, B. L., Solomon, J., Brock, J. F. 1971. The effect of hypophysectomy and growth hormone replacement on albumin synthesis and catabolism in the rat. *Biochem. J.* 24:529–35
91. Reaven, E. P., Peterson, D. T., Reaven, G. M. 1973. The effect of experimental diabetes mellitus and insulin replace-

ment on hepatic ultrastructure and protein synthesis. *J. Clin. Invest.* 52:248–62

92. Baligno, H. F., Neuhaus, O. W. 1970. Effect of insulin on the injury stimulated synthesis of serum albumin in the rat. *Life Sci.* 9(II):1039–44
93. Bjφrneboe, M. 1946. Studies on the serum proteins in hepatitis. I. The relation between serum albumin and serum globulin. *Acta Med. Scand.* 123:393–401
94. Rothschild, M. A., Oratz, M., Franklin, E. C., Schreiber, S. S. 1962. The effect of Hypergammaglobulinemia on albumin metabolism in hyperimmunized rabbits studied with albumin I^{131}. *J. Clin. Invest.* 41:1564–71
95. Oratz, M. 1970. Oncotic pressure and albumin synthesis. In *Plasma Protein Metabolism,* ed. M. A. Rothschild, T. Waldmann, 223–38. New York: Academic. 560 pp.
96. Rothschild, M. A., Oratz, M., Evans, C., Schreiber, S. S. 1964. Alterations in albumin metabolism after serum and albumin infusions. *J. Clin. Invest.* 43: 1874–80
97. Dich, J., Hansen, S. E., Thieden, H. D. 1973. Effect of albumin concentration and colloid osmotic pressure on albumin synthesis in the perfused rat liver. *Acta Physiol. Scand.* 89:352–58
98. Goldstein, L., Forster, R. P. 1971. Osmoregulation and urea metabolism in the little skate raja erinacea. *Am. J. Physiol.* 220:742–46
99. Rasmussen, L. E., Rasmussen, R. A. 1967. Comparative protein and enzyme profiles of the cerebrospinal fluid, extradural fluid, nervous tissue, and sera of elasmobranchs. In *Sharks, Skates and Rays,* ed. P. W. Gilbert, R. F. Mathewson, D. P. Rall, 361–80. Baltimore: John Hopkins. 624 pp.
100. Oratz, M., Walker, C., Schreiber, S. S., Rothschild, M. A. 1967. Albumin and fibrinogen metabolism in heat-and cold-stressed rabbits. *Am. J. Physiol.* 213: 1341–49
101. Surks, M. I. 1966. Metabolism of human serum albumin in man during acute exposure to high altitude (14,100 feet). *J. Clin. Invest.* 45:1442–51
102. Schnakenberg, D. D., Krabill, L. F., Weiser, P. C. 1971. The anorexic effect of high altitude on weight gain, nitrogen retention and body composition of rats. *J. Nutr.* 101:787–96
103. Post, J., Patek, A. J. Jr. 1942. Serum proteins in cirrhosis of the liver. *Arch. Intern. Med.* 69:67–82, 83–89
104. Hasch, E., Jarnum, S., Tygstrup, N. 1967. Albumin synthesis rate as a measure of liver function in patients with cirrhosis. *Acta Med. Scand.* 182:83–92
105. Dykes, P. W. 1961. A study of the effects of albumin infusions in patients with cirrhosis of the liver. *Quart. J. Med.* 30:297–327
106. Wilkinson, P., Mendenhall, C. L. 1963. Serum albumin turnover in normal subjects and patients with cirrhosis measured by ^{131}I-labeled human albumin. *Clin. Sci.* 25:281–92
107. Dykes, P. W. 1968. The rates of distribution and catabolism of albumin in normal subjects and in patients with cirrhosis of the liver. *Clin. Sci.* 34: 161–83
108. Rothschild, M. A. et al 1969. Albumin synthesis in cirrhotic subjects with ascites studied with carbonate-^{14}C. *J. Clin. Invest.* 48:344–50
109. Cain, G. D., Maejer, G., Jones, E. A. 1970. Augmentation albumin but not fibrinogen synthesis by corticosteroids in patients with hepatocellular disease. *J. Clin. Invest.* 49:2198–204
110. Westergaard, H., Jarnum, S., Ramsoe, K., Ranek, L. 1972. Albumin and immunoglobulin turnover in patients with chronic active liver disease treated with prednisone. *Scand. J. Gastroenterol.* 7: 623–30
111. Waldmann, T. A., Strober, W. 1969. Metabolism of immunoglobulins. *Progr. Allergy* 13:1–110
112. Wilkinson, P., Pinto, B., Senior, J. R. 1965. Reversible protein-losing enteropathy with intestinal lymphangiectasia secondary to chronic constrictive pericarditis. *N. Engl. J. Med.* 273: 1178–81
113. Jeejeebhoy, K. N. et al 1969. Metabolism of albumin and fibrinogen in patients with tropical sprue. *Gastroenterology* 56:252–67
114. Jones, E. A., Craigie, A., Tavill, A. S. 1968. Protein metabolism in the intestinal stagnant loop syndrome. *Gut* 9: 466–69
115. Brooks, S. G., Dobbins, W. O. III. 1972. Autoradiographic localization of I^{125}-labeled albumin in the intestine of guinea pigs: A light and electron microscopic study. *Gastroenterology* 62: 1001–11
116. Wochner, R. D., Weisman, S. M., Waldmann, T. A., Houston, D., Berlin, N. I. 1968. Direct measurement of the rates of synthesis of plasma proteins in control subjects and patients with gas-

trointestinal protein loss. *J. Clin. Invest.* 47:971–82

117. Bianchi, R. et al 1969. Patterns of serum albumin metabolism in renal diseases. In *Physiology and Pathophysiology of Plasma Protein Metabolism,* ed. G. Birke, R. Norberg, L. O. Plantin, 195–211. London: Pergamon
118. Blahd, W. H., Fields, M., Goldman, R. 1955. The turnover rate of serum albumin in the nephrotic syndrome as determined by I^{131}-labelled albumin. *J. Lab. Clin. Med.* 46:747–56
119. Jensen, H. 1969. Plasma protein metabolism in the nephrotic syndrome in adults. See Ref. 117, pp. 213–19
120. Jensen, H., Rossing, N., Andersen, S. B. 1967. Albumin metabolism in the nephrotic syndrome in adults. *Clin. Sci.* 33:445–57
121. Mogielnicki, R. P., Waldmann, T. A., Strober, W. 1971. Renal handling of low molecular weight proteins. I. L-chain metabolism in experimental renal disease. *J. Clin. invest.* 50:901–9
122. Katz, J., Bonorris, G., Sellers, A. L. 1964. Effect of nephrectomy on plasma albumin catabolism in experimental nephrosis. *J. Lab. Clin. Med.* 63:680–86
123. Marsh, J. B., Drabkin, D. L. 1958. Metabolic channeling in experimental nephrosis. II. Influence of diet and of adrenalectomy; liver hypertrophy. *J. Biol. Chem.* 230:1063–71; Ibid. IV. Net synthesis of plasma albumin by liver slices from normal and nephrotic rats. 230:1073–81
124. Smuckler, E. A., Benditt, E. P. 1965. Studies on carbon tetrachloride intoxication. III. A subcellular defect in protein synthesis. *Biochemistry.* 4:671–79
125. Smuckler, E. A., Parthier, B., Hultin, T. 1968. The effects of polyuridylic acid on phenylalanine incorporation by subcellular fractions from carbon tetrachloride-poisoned rat liver. *Biochem. J.* 107:151–63
126. Smuckler, E. A., Iseri, O. A., Benditt, E. P. 1962. An intracellular defect in protein synthesis induced by carbon tetrachloride. *J. Exp. Med.* 116:55–72
127. Rubin, E., Lieber, C. S. 1968. Alcohol-induced hepatic injury in nonalcoholic volunteers. *N. Engl. J. Med.* 278: 869–76
128. Rochschild, M. A., Oratz, M., Schreiber, S. S. 1972. Effects of carbon tetrachloride on albumin synthesis. *J. Clin. Invest.* 50:2310–14
129. Rothschild, M. A., Oratz, M., Mongelli, J., Schreiber, S. S. 1972. Alcohol induced depression of albumin synthesis: Reversal by tryptophan. *J. Clin. Invest.* 51:1812–18
130 Jeejeebhoy, KLN., Phillips, M. J., Bruce-Robertson, A., Ho, J., Sodtke, U. 1972. The acute effect of ethanol on albumin, fibrinogen and transferrin synthesis in the rat. *Biochem. J.* 126: 1111–26
131 Regoeczi, E., Wong, K. L. 1972. Effects of intravenous neuraminidase on the turnover of fibrinogen. In *Protein Turnover Ciba Foundation Symposium* 9 (New Series): 181–201
132. Morell, A. G., Gregoriadis, G., Scheinberg, I. H., Hickman, J., Ashwell, G. 1971. The role of sialic acid in determining the survival of glycoproteins in the circulation. *J. Biol. Chem.* 246:1461–67
133. McFarlane, A. S., Koj, A. 1970. Short term measurement of catabolic rates using iodine labeled plasma proteins. *J. Clin. Invest.* 49:1903–11
134. Cohen, S., Gordon, A. H. 1958. Catabolism of plasma albumin by the perfused rat liver. *Biochem. J.* 70:544–51
135. Freeman, T., Gordon, A. H., Humphrey, J. H. 1958. Distinction between catabolism of native and denatured proteins by the isolated perfused liver after carbon loading. *Brit. J. Exp. Pathol.* 39:459–71
136. Beeken, W. L. 1965. In vitro metabolism of human serum albumin by subcellular fractions of hepatic tissue from man. *J. Lab. Clin. Med.* 65:649–58
137. Gabathuler, M. P., Rhyser, J. H. P. 1969. The digestive function of lysosomes as studied by the turnover of ingested foreign macromolecules. *Proc. Roy. Soc. B* 173:95–98
138. Ehrenreich, B. A., Cohn, Z. A. 1968. Fate of hemoglobulin pinocytosed by macrophages in vitro. *J. Cell Biol.* 38:244–48
139. Goldfischer, S. et al 1970. Hemoglobin uptake by rat hepatocytes and its breakdown within lysosomes. *J. Cell Biol.* 44:513–29
140. Kerr, R. M., Dubois, J. J., Holt, P. R. 1967. Use of ^{125}I- and ^{51}Cr-labeled albumin for the measurement of gastrointestinal and total albumin catabolism. *J. Clin. Invest.* 46:2064–82
141. Hardwicke, J., Rankin, J. G., Baker, K. J., Preisig, R. 1964. The loss of protein in human and canine hepatic bile. *Clin. Sci.* 26:509–17
142. Ancell, H. 1839–40. Course of lectures on the physiology and pathology of the blood and the other animal fluids. *Lancet* 1:222 (As quoted from Serum albumin. 1969. *Am. J. Dig. Dis.* 14:711–44)

Copyright 1975. All rights reserved

LEFT VENTRICULAR FUNCTION IN ACUTE AND CHRONIC CORONARY ARTERY DISEASE[1]

❖7135

Charles E. Rackley, M.D. and Richard O. Russell, Jr., M.D.
Division of Cardiology, Department of Medicine, University of Alabama Medical Center,
Birmingham, Alabama 35294

INTRODUCTION

The understanding and evaluation of left ventricular function in patients with acute or chronic disturbances caused by coronary artery disease have significantly contributed to patient management. Such information can be utilized in the immediate care and management of the patient as well as in the selection of medical or surgical forms of treatment. Progress has been made in the modification of invasive techniques for application in coronary care units. Improved methods for the analysis of left ventricular function have also been utilized in diagnostic cardiac catheterization laboratories for patients with chronic coronary artery disease. This review will present the methods for evaluating left ventricular function in patients with acute myocardial infarction and the techniques for assessing ventricular performance in the chronic stage of coronary artery disease.

METHODS

Patients with Acute Myocardial Infarction

Present techniques for the acquisition of information on the heart and circulatory system in patients with acute myocardial infarction involve right and left heart

---

[1] This research was supported in part by the Myocardial Infarction Research Unit Program, Contract Number HL 43-67-1441, the Cardiovascular Research and Training Center, Program Project Grant Number HL 11,310 (Division of Heart and Vascular Diseases, National Heart and Lung Institute), and the Clinical Research Unit Grant Number RR-32 (General Clinical Research Centers Program, Division of Research Resources, National Institutes of Health).

catheterization. In the right heart procedures an antecubital area is selected for local anesthesia and a catheter introduced into a vein with advancement to the pulmonary artery under fluoroscopic visualization or with pressure monitoring (1, 2). In the past, standard cardiac catheters have been employed for this approach but the developmemt of the balloon-tipped Swan-Ganz catheter has enabled the right heart procedure to be performed without fluoroscopic control (3). The Swan-Ganz catheter can be advanced to the pulmonary artery with pressure monitoring and the balloon inflated to provide pulmonary capillary or wedge pressure. With deflation of the balloon, pulmonary artery systolic and diastolic pressure as well as right atrial pressure or central venous pressure can be recorded from the two lumina in the catheter. In left heart studies a brachial artery or femoral artery is entered and the catheter advanced retrograde into the left ventricle. Under fluoroscopic visualization the catheter can be directed into the left ventricle or the percutaneous femoral technique can be employed with a special preformed catheter for pressure monitoring (4). In patients studied with cardiogenic shock or extreme hypotension a cutdown of the brachial artery may be necessary to introduce the catheter. Therefore, with the employment of these techniques pressures can be recorded from the left ventricle, pulmonary artery, and right atrium.

In the coronary care unit, cardiac output can be measured by the standard green dye indicator dilution technique, the thermodilution method, or in certain patients with echocardiography (1, 5). For the green dye technique the indicator is injected into the pulmonary artery and simultaneous withdrawal obtained from a systemic artery. Although the thermodilution method employs the same basic assumptions used in the green dye technique, the incorporation of a thermistor into the Swan-Ganz catheter permits injection of the cold solution into the right atrium and temperature sampling in the pulmonary artery (5). In patients in whom high quality echocardiograms can be obtained in the coronary care unit, measurements of end-diastolic and end-systolic volume can be obtained and the cardiac output calculated from the left ventricular stroke volume and the heart rate. In addition to preliminary measurements of intra-cardiac pressures along with cardiac output, the catheters in the pulmonary artery and systemic arteries can remain for periods of days with constant recording of chamber and vascular pressures (7). The cardiac output can be measured at intervals and during changes in clinical condition or during therapeutic interventions.

During the period of hemodynamic monitoring, interventions can be utilized to evaluate ventricular function. The blood volume can be expanded with a plasma expander, and the alterations in left ventricular filling pressure and cardiac output recorded (8). Sequential measurements of left ventricular filling pressure and cardiac index with the derived parameters of stroke index, stroke work index, and stroke power index can be related to describe a ventricular function curve. In patients with a left ventricular catheter the rate of development of isovolumic pressure can be analyzed at rest and after inotropic interventions for an additional measure of muscle performance (9). Finally, these measurements of filling pressure and cardiac output can be obtained before and after pharmacologic interventions with agents such as diuretics, digitalis, or vasoactive substances (10–12). Therefore, with the appropriate facilities, measurements of pressure, cardiac output, derived indices, and ventricular performance can be assessed in a serial manner in patients with acute myocardial infarction.

Patients with Chronic Coronary Artery Disease

The basic techniques for measurement of right and left heart pressures and cardiac output are employed in patients with stable coronary artery disease in the diagnostic cardiac catheterization laboratory. A clinical indication for these studies is the delineation of coronary artery anatomy and assessment of left ventricular function. The techniques currently employ the Sones' approach from the brachial artery or the Judkins' method from the femoral artery (13, 14). In addition to coronary angiography, information on the left ventricular performance can be obtained from measurements and analysis of chamber pressure which include the left ventricular end-diastolic pressure, A-wave amplitude, dp/dt, and V_{max} (15–17). Ventriculography is performed with biplane angiocardiography and single or biplane cineangiography. Measurements of left ventricular chamber dimensions and wall thickness can be obtained to calculate the end-diastolic volume, end-systolic volume, left ventricular stroke volume, ejection fraction, and regurgitant volume. The method of Dodge and co-workers is employed to calculate chamber dimensions from biplane angiocardiograms. The method of Greene and associates is often used to estimate ventricular dimensions from single plane cineangiograms and the method of Rackley and associates to determine biplane cineangiographic measurements (18–20). The thickness of the lateral left ventricular wall can be added to the chamber dimensions and the mass of the left ventricle calculated (21). Chamber dimensions, wall thickness, and chamber pressures can be incorporated in equations to calculate wall forces or stress throughout the cardiac cycle (22). Compliance of the left ventricle can be calculated from diastolic left ventricular pressure and single plane cineangiograms or biplane angiocardiograms (23, 24). Several methods are used for the identification, analysis, and quantitation of abnormalities of left ventricular wall motion in patients with chronic coronary artery disease (25–27).

In addition to the resting parameters of left ventricular function that can be obtained in these patients, various provocative tests have been applied to examine cardiac function under conditions of increased left ventricular performance. Exercise studies can be performed with either active leg movement or bicycle ergometry in the supine position (28–30). Atrial pacing has also been utilized to assess changes in left ventricular performance at increasing heart rates (31–34). The expansion of the blood volume has been employed to construct ventricular function curves (35–37). Pharmacologic agents which alter afterload or the contractile state of the left ventricle can be administered with the recording of these measurements before and after intervention (38, 39). Left ventricular angiography can be repeated during pacing or after drug administration to reassess wall motion both in abnormal areas and in areas of normal movement (40–44).

RESULTS

Patients with Acute Myocardial Infarction

The earliest reported hemodynamic measurements in patients with acute myocardial infarction included observations of central venous pressure and cardiac output (45–47). These investigators described a spectrum of findings in patients presenting with mild myocardial infarctions to those presenting with left ventricular failure or cardiogenic shock. In the mild clinical presentations, central venous pressure was normal and cardiac output was calculated within the normal range. Generally, the central venous pressure was abnormally elevated and the cardiac output reduced in patients with heart failure. Patients in cardiogenic shock often presented the highest central venous pressure and the lowest cardiac output. Occasionally, the central venous pressure was normal in patients who developed cardiogenic shock (48–51).

Various explanations have been proposed to explain the mechanism for the elevation of central venous pressure in patients with acute myocardial infarction. Nixon and associates (52) suggested predominant peripheral venoconstriction in the elevation of central venous pressure, but Cohn and his colleagues (53) considered a relationship between central venous pressure and left ventricular filling pressure. However, a comparison between central venous pressure and left ventricular end-diastolic pressure in patients with acute myocardial infarction failed to reveal a significant correlation (54). Rapaport & Scheinman demonstrated a poor correlation between central venous pressure and pulmonary artery end-diastolic pressure in clinical myocardial infarction (55). Based on these clinical studies, central venous pressure is not considered a reliable estimate of left ventricular filling pressure in acute infarction (56, 57). Rarely, patients with inferior myocardial infarction may present with abnormal elevations of central venous pressure due to extension of the left ventricular infarction into the free wall of the right ventricle (58). Therefore, an elevated central venous pressure indicates abnormalities in right ventricular diastolic filling and function in myocardial infarction, but the measurement does not accurately reflect acute disturbances in left ventricular performance (59).

Subsequent investigators performed right heart catheterizations with acute myocardial infarction. Rutherford and co-workers reported abnormal elevations in the mean pulmonary artery pressure in 16 of 25 patients which was associated with a complicated clinical course (60). In addition to the measurements of mean pulmonary artery pressure and wedge or capillary pressure, the pulmonary artery end-diastolic pressure has been utilized to estimate the left ventricular end-diastolic or filling pressure in patients with acute myocardial infarction. Investigators reported a linear relationship between pulmonary artery end-diastolic and left ventricular end-diastolic pressure ranging from 5 to 20 mm Hg in patients with congenital or acquired heart disease and no evidence of mitral valve obstruction or severe pulmonary disease (61). Subsequent reports questioned the relationship between pulmonary artery and left ventricular end-diastolic pressure in chronic disease states and discrepancies were described between these measurements. Several investigators have reported disparities between pre and post A-wave left ventricular end-diastolic pressure and pulmonary artery end-diastolic pressure in the chronic disease state

(62, 63). In patients with acute myocardial infarction small differences between the pulmonary artery end-diastolic and left ventricular end-diastolic pressure have been analyzed (64). However, other workers have reported similar values for pulmonary artery end-diastolic and left ventricular end-diastolic pressure in patients with acute myocardial infarction (65). Even in patients with cardiogenic shock and abnormalities of hypoxia, acidosis, and variations of heart rate, a consistent relationship was shown between pulmonary artery end-diastolic and left ventricular end-diastolic pressure (66). Thus, there appears to be sufficient clinical evidence to support the use of pulmonary artery end-diastolic pressure as an estimate of left ventricular end-diastolic pressure in patients with acute myocardial infarction when there is no accompanying mitral valve disease or severe pulmonary dysfunction.

The measurement of left ventricular filling pressure as pulmonary capillary or pulmonary end-diastolic pressure has revealed a high incidence of abnormal values in patients with acute myocardial infarction. In the absence of the clinical signs of left ventricular failure, the pulmonary artery end-diastolic pressure is frequently abnormal. In one series, 14 out of 21 patients (67%) considered clinically uncomplicated revealed an initial pulmonary artery end-diastolic pressure above 12 mm Hg (67). Similarly, Rotman and associates described values for pulmonary artery end-diastolic pressure greater than 14 mm Hg in 16 of 34 patients (47%) without clinical signs of left ventricular failure (68). In patients with mild to severe left ventricular failure, progressive elevations in left ventricular filling pressure indicate severity of depressed left ventricular function. Although severe left ventricular failure had been assumed to accompany cardiogenic shock in patients with acute myocardial infarction, Nixon and associates first reported a normal left atrial pressure in a patient with cardiogenic shock (52). Additional investigations have confirmed this finding but a wide range of left ventricular filling pressures from normal to extreme elevations have been reported (69, 70). The measurement of left ventricular filling pressure in patients with acute myocardial infarction provides a very sensitive expression for impaired left ventricular function in acute myocardial infarction.

In addition to the measurements of right and left heart pressures in patients with acute myocardial infarction, the determination of cardiac output has continued to contribute to the assessment of left ventricular function. Early studies reported normal values for resting cardiac index in patients with mild or clinically uncomplicated infarctions but the cardiac index may be significantly reduced in 50% of these patients (67). Patients with anterior myocardial infarction by the electrocardiogram present greater depression of stroke index than with inferior infarction (65). Patients presenting with severe heart failure or cardiogenic shock generally exhibit significant depressions of cardiac output. Although the methods for measuring cardiac output become less reliable at low flow rates, the lowest recorded values for cardiac output usually occur in severe cardiac shock. Occasionally, the measurement of cardiac output in a patient with clinical signs of left ventricular failure may be normal. Analysis of the circulatory conditions will reveal that this is not a valid resting output determination, since the compensatory mechanisms for maintaining cardiac output may be maximally stimulated. The hemodynamic evaluation of left ventricular performance can also contribute to the immediate prognosis. Although the mortality for patients presenting in cardiogenic shock exceeds 80%, the initial measurements of left ventricular filling pressure and cardiac index have enabled four subgroups to be identified (71). Patients presenting a pulmonary artery end-diastolic or left ventricular end-diastolic pressure greater than 29 mm Hg did not survive in any instance. In patients with the left ventricular filling pressure greater than 15 mm Hg and the cardiac index less than 2.0 L/min/m^2, the mortality was 92%. A third subgroup presented an initial left ventricular filling pressure less than 15 mm Hg with an index less than 2.0 L/min/m^2 and the mortality was 63%. A fourth group presented a left venticular filling pressure less than 29 mm Hg but a cardiac index greater than 2.0 L/min/m^2 and the mortality was 13%. Therefore, initial hemodynamic measurements in patients with acute infarction and cardiogenic shock can predict prognosis in subgroups with mortalities of 100%, 92%, 63%, and 13%.

The measurements of left ventricular filling pressure and cardiac output have been examined in patients with acute infarction following the expansion of blood volume with agents such as low molecular weight dextran. Several investigators have reported that rapid volume expansion in these patients and elevations in left ventricular filling pressure are associated with increases in the cardiac index and derived parameters of systolic ventricular performance (1,8,51). The changes induced in left ventricular function following volume expansion may persist for 6-10 hr in patients with acute myocardial infarction. The serial changes in left ventricular filling pressure and cardiac output during the rapid infusion of dextran can be incorporated to construct a left ventricular function curve. These function curves are characterized by the initial slope, the plateau, and in some patients a descending limb (8,72). The ventricular function curves may be depressed or flat during the first 24 hr after acute myocardial infarction (73). However, subsequent function curves obtained in this manner on the second or third day after the acute infarction often reveal a spontaneous fall in left ventricular filling pressure and a rise in resting cardiac output (74). Following the volume expansion on the second or third day after acute infarction, the function curve may shift upwards and to the left and be characterized by a steeper slope with the loss of the descending limb (1). Left ventricular function curves can be constructed using the relationship of stroke index, stroke work index, or stroke power index to the left ventricular filling pressure.

Left ventricular function has been further examined in patients with acute myocardial infarction by the analysis of left ventricular end-diastolic pressure, contractile element velocity, and maximum dp/dt (9). Parmley and co-workers reported considerable overlap between survivors and nonsurvivors in the levels of contractile element velocity, left ventricular end-diastolic pressure, cardiac index, heart rate, and systemic vascular resistance. There was fairly good separation of survivors and nonsurvivors with measurements of maximum left ventricular dp/dt, mean isovolumic $\Delta p/\Delta t$, mean arterial pressure, and stroke work. However, the changes in contractile element velocity were no more sensitive than ventricular function curves in detecting changes in the contractile state. These investigators concluded that measurement of left ventricular pressure would provide little more information of clinical value than could be obtained from standard techniques employing the pulmonary artery balloon catheter and measurements of cardiac output.

The assessment of left ventricular function in acute infarction can be further studied after the administration of diuretics and inotropic agents (10–12). Diuretic agents produce a fall in the left ventricular filling pressure, and this can occur before the onset of diuresis. The reduction in left ventricular filling pressure and diuresis may be associated with no measurable change or a slight decrease in cardiac output (10). Digitalis has been shown to exert the most significant hemodynamic changes in acute pulmonary edema with a resultant fall in left ventricular filling pressure and an augmentation of the cardiac index (75). In patients with absent or minimal heart failure as well as cardiogenic shock, digitalis may not significantly alter the left ventricular filling pressure or the cardiac output (11,71). Propranolol has produced reductions in heart rate and cardiac index without changes in left ventricular filling pressure (12). Blood pressure lowering agents such as nitroprusside or nitroglycerin can reduce filling pressure and increase cardiac index in patients with severe heart failure and pulmonary edema (76,77). The dextran infusion left ventricular function curves have revealed that the optimal left ventricular filling pressure in acute myocardial infarction ranges from 20 to 24 mm Hg. Therefore, volume expansion or reduction can be clinically controlled to provide the optimal range for the filling pressure during the initial phases of convalescence.

The importance of serially evaluating left ventricular function in acute myocardial infarction has been demonstrated by several groups of investigators (1,7,72,78–80). The initial left ventricular filling pressure, which is often abnormally elevated, has been observed to decline during the initial convalescent period. Similarly, the reduced cardiac output has been documented to rise from the initially reduced values when measured over a span of several days (1,7,72). These hemodynamic observations suggest spontaneous improvement in left ventricular function following the acute infarction. Ventricular function curves obtained serially during the first three days after infarction have exhibited steeper slopes, absence of the descending limb on the second or third day, and a shift upward and to the left, indicating improvement in mechanical performance (18,8,72). Early changes in pulmonary artery end-diastolic pressure and echocardiographic left ventricular dimensions have been related to assess left ventricular compliance after acute infarction (81). The sequential changes in filling pressure, echocardiographic dimensions, and left ventricular compliance demonstrated a close correlation with the clinical course and survival in patients observed for eight days after acute myocardial infarction. Finally, continuous hemodynamic monitoring has provided objective physiologic information on the pharmacologic response in patients with heart failure and cardiogenic shock.

Patients with Chronic Coronary Artery Disease

RESTING HEMODYNAMIC MEASUREMENTS Several investigators have reported measurements of left ventricular end-diastolic pressure, cardiac output, and derivatives of these measurements in patients with stable coronary disease manifest by angina pectoris or a previous myocardial infarction (23,29,30,32,34,82–84). Normal and abnormal values for left ventricular end-diastolic pressure and cardiac output at rest have been reported in patients with stable coronary artery disease. Generally, patients with a history of angina pectoris have normal resting hemodynamics (29,30,32,34). Hamby and co-workers observed no significant abnormalities of left ventricular end-diastolic pressure in patients with isolated disease of the right coronary artery but they did find abnormal measurements in single disease of the left anterior descending coronary artery and in diffuse vessel disease (83). In the comparison of patients with one, two, and three vessel coronary artery disease either with or without a previous history of myocardial infarction, Moraski and colleagues encountered statistically significant abnormalities of resting left ventricular end-diastolic pressure only in patients with three vessel involvement and a previous myocardial infarction (84). In patients studied within 12 months after myocardial infarction, Smith and co-workers documented a left ventricular end-diastolic pressure above 12 mm Hg in 22 of 42 patients (16). These investigators further observed that the amplitude of the left ventricular A-wave was related to the end-diastolic pressure. These investigations indicate that resting hemodynamic measurements are frequently normal in patients with coronary artery disease, but abnormalities can be observed in patients with disease of the left anterior descending coronary, diffuse coronary involvement, and after a myocardial infarction.

Investigators have calculated left ventricular volume in patients with coronary artery disease and related these measurements to the extent of the anatomical coronary artery disease. Bristow and associates recorded normal values for end-diastolic volume in 15 patients with coronary artery disease but observed an end-diastolic pressure greater than 12 mm Hg in 8 patients (23). In patients with severe coronary artery disease, Rackley and co-workers documented significant mitral regurgitation in 15 of 23 patients with an end-diastolic volume greater than 154 ml/m^2 (82). In a study of single vessel coronary artery disease, Hamby et al reported an end-diastolic volume larger than 100 ml/m^2 only in patients with disease in the left anterior descending coronary artery (83). Hamilton et al (26) and Falsetti and colleagues (85) described a significant increase in end-diastolic volume in all patients with ischemic heart disease and a history of heart failure. Feild and co-workers observed an abnormal increase in end-diastolic volume in 11 of 25 patients studied within one year after acute infarction. Additionally, Feild and associates estimated the size of the abnormally contracting segment in the left ventricle and concluded that an area greater than 17% produced dilatation (27). Moraski et al calculated that statistically significant increases in end-diastolic volume occurred with one, two, and three vessel disease with a previous myocardial infarction (84).

The determination of left ventricular mass in patients with stable coronary artery disease has revealed a relationship between abnormal increases in left ventricular end-diastolic volume or dilatation and left ventricular mass or hypertrophy (26, 82,84,86). As the end-diastolic volume increases as a chronic compensatory mechanism to maintain stroke volume, secondary hypertrophy or an increase in left ventricular mass has also been observed. Hamilton et al recorded abnormal values for left ventricular mass in patients with ischemic heart disease and contraction abnormalities both with and without a history of infarction (26). Moraski reported a significant increase in left ventricular mass in patients with three vessel coronary disease with and without a previous infarction (84). Gould and co-workers documented left ventricular hypertrophy after myocardial infarction in proportion to left ventricular dilatation in coronary artery disease and further described a syndrome of massive hypertrophy, hypokinesis, and congestive heart failure which was quantitatively identical to that found in primary cardiomyopathy (86).

Several methods of data collection and analysis have been utilized to calculate compliance of the left ventricle in patients with coronary artery disease (23,24,87). Diamond & Forrester calculated from left ventricular pressure a measurable change in ventricular compliance with the development of coronary artery disease and a further increase in wall stiffness with an acute infarction (87). Bristow et al documented abnormalities of diastolic compliance in patients with coronary artery disease without significant increase in end-diastolic volume (23). Smith and co-workers computed compliance values by four formulae, and the data suggested that left ventricular compliance was sensitive to the magnitude of segmental disease and influenced less by changes in ventricular size in postmyocardial infarction patients (24). In patients with only a history of angina pectoris and disease in either one, two, or three coronary arteries, compliance values at rest have remained within normal limits (84).

Ventricular function in stable coronary artery disease has been assessed by methods which evaluate the mechanical or pump performance of the ventricle and other methods which assess the state of the muscle or myocardium. The ejection fraction is a frequently employed expression of total mechanical performance of the left ventricle and may be abnormally reduced in patients with either one, two, or three vessel coronary artery involvement and a previous history of myocardial infarction (84). In addition, the ejection fraction can be significantly depressed in patients with three vessel coronary artery disease without a previous clinical history of myocardial infarction. In chronic coronary artery disease, patients with congestive heart failure exhibited ejection fractions less than 0.40. Also, an audible left ventricular gallop sound is associated with an ejection fraction less than 0.30 (26,88). The assessment of muscle performance in patients with coronary disease has revealed abnormal values for $V_{\max}$ and V_{pm} [$V_{\mathrm{pm}} = [(\mathrm{d}p/\mathrm{d}t/32P]_{\max}$], and these abnormalities have been demonstrated to correlate with the ejection fraction (89).

The assessment of wall motion in patients with stable coronary artery disease as well as those with previous myocardial infarction has been investigated in recent years. A commonly used terminology employs akinesis for segments of the ventricular wall that exhibit no motion from diastole through systole, hypokinesis for those segments of the wall that exhibit reduced motion, and dyskinesis for those areas of the ventricular myocardium that expand during systole (90). In addition to the descriptive nomenclature for abnormal wall motion in patients with coronary artery disease, methods have been developed for quantitating the extent of the abnormal wall motion as well as for analyzing segments of the ventricular myocardium throughout the cardiac cycle. Feild and co-workers described a method for superimposing diastolic and systolic biplane angiocardiograms so that areas of abnormal wall motion could be quantitated and expressed as a percent abnormally contracting segment for the entire ventricle (27). Herman & Gorlin divided the total ventriculogram into six or more segments so that each axis and segment can be individually analyzed and quantitated (90). In patients studied 2 to 12 months after acute myocardial infarction, the size of the abnormally contracting segment is assumed to be the residual scar from the myocardial infarction. The size of the abnormally contracting segment is related to abnormalities in compliance, end-diastolic pressure, end-diastolic volume, and the clinical manifestations of heart failure. An abnormally contracting segment greater than 8% in the chronic stages after acute infarction is associated with abnormal values of compliance (24). A segment greater than 15% is associated with abnormal elevations in left ventricular end-diastolic pressure (27). An abnormally contracting segment greater than 17% results in dilatation or abnormal increases in end-diastolic volume (27). Segments greater than 23% are associated with clinical manifestations of left ventricular failure and a high incidence of audible ventricular gallop sounds (27).

PROVOCATIVE HEMODYNAMIC MEASUREMENTS Physical exercise has been a traditional method of assessing cardiac performance, but only limited investigations have been performed with simultaneous measurement of left ventricular pressure. Foster & Reeves reported significant reduction of cardiac output, stroke volume, and mean rate of systolic ejection in patients with angina pectoris as compared to comparable middle-aged or younger groups of normals (91). Cohen et al compared exercise hemodynamics in patients who developed angina to those without pain and concluded that with angina the left ventricular end-diastolic pressure did not uniformly rise but the response of the cardiac index was subnormal (29). Parker and colleagues studied subjects with a previous history of angina and normal subjects and reported normal left ventricular end-diastolic pressures and cardiac indices in the patients at rest (30). With exercise, the anginal patients exhibited an abnormal increase in the left ventricular end-diastolic pressure and subnormal increases in cardiac index. These exercise studies would suggest that depressed left ventricular function can occur during exercise in patients with angina pectoris.

A recent method of increasing the demands on cardiac performance has been atrial pacing with measurements of left ventricular pressure and cardiac output (31–34). Parker and associates induced angina with atrial pacing and initially observed a fall in left ventricular end-diastolic pressure in normal subjects as well as in patients who developed angina (33). However, with continued pacing the left ventricular end-diastolic pressure rose in the patients with chest pain. O'Brien and colleagues compared hemodynamic findings in angina produced by atrial pacing and exercise with the observation that left ventricular end-diastolic pressure was normal with pacing-induced angina and abnormally elevated with exercise-provoked angina (32). Parker et al obtained similar results and also noticed that following the cessation of pacing the left ventricular end-diastolic pressure rose abnormally in patients who had experienced angina (33). Parker proposed that the ischemic ventricular function curve and the depression was reversible after pacing. In patients with coronary artery disease who failed to develop pain during pacing, Linhart and co-workers obtained normal hemodynamic measurements and concluded that patients developing angina with pacing may or may not have abnormal hemodynamic findings (39).

Ventricular performance in stable patients has been evaluated by various interventions which permit the construction of ventricular function curves. These curves have been produced by the expansion of the blood volume with dextran infusion, changes associated with the infusion or pressor agents, changes induced with atrial pacing, and by the hemodynamic alterations following the injection of radiographic contrast material (33,35–39). In patients with stable angina the infusion of dextran has reduced the threshold for pacing-induced angina (35). The dextran infusion ventricular function curves in patients with a previous history of myocardial infarction have demonstrated normal ascending slopes, depressed function curves, and flat curves (36,92). The shape of the function curve has demonstrated a significant relationship to the ejection fraction. In patients receiving angiotensin for the elevation of arterial blood pressure, both normal and abnormal function curves have been described in patients with coronary artery disease (39). In patients with abnormally elevated resting left ventricular end-diastolic pressure, angiotensin function curves have all been abnormal. In patients with abnormal angiocardiograms, angiotensin function curves are similarly abnormal. The pacing ventricular function curves reveal both normal and abnormal responses in patients with coronary artery disease (33). The ventricular function curves constructed after angiographic contrast injection are similar to those obtained after dextran infusion (39). These findings are similar to those patients that have volume expansion after angiographic contrast studies.

In addition to the hemodynamic measurements and parameters of ventricular function analysis, wall motion in cineangiograms has been examined during pacing and drug infusion studies (40–44). With increase in heart rate from atrial pacing, increased abnormalities in wall motion have been described as well as further changes in left ventricular compliance (40). The administration of nitroglycerin has shown a reduction in the size of the abnormal wall motion as well as enhanced movement in areas of diminished wall motion observed under resting conditions (42,43). Epinephrine has also been infused with repeat angiograms, and wall motion abnormalities have been observed to diminish (44). Following either nitroglycerin administration or epinephrine infusion, the hypokinetic areas and normally contracting myocardium have been shown to increase the extent of wall motion.

SUMMARY

The evaluation of left ventricular function in patients with acute myocardial infarction has shown:

1. Limitations in the use and interpretation of central venous pressure.
2. Pulmonary artery end-diastolic pressure reflects left ventricular end-diastolic pressure in the absence of pulmonary vascular or mitral valve disease.
3. Frequent elevations of left ventricular filling pressure in mild or clinically uncomplicated infarction.

4. Anterior infarctions present greater depression of left ventricular function than inferior infarctions.
5. Initial hemodynamic measurements in cardiogenic shock can predict prognosis with medical management.
6. Left ventricular function frequently improves during the early convalescent period.
7. Hemodynamic monitoring can be useful in following changes in left ventricular function and the response to therapy.

The assessment of left ventricular performance in patients with chronic heart disease has shown:

1. Resting hemodynamic measurements are often normal but abnormalities can be observed in patients with disease of the left anterior descending coronary artery, diffuse coronary involvement, and after myocardial infarction.
2. Increases in end-diastolic volume or dilatation and left ventricular mass or hypertrophy can develop in severe coronary disease and after myocardial infarction.
3. The size of abnormally contracting segment after myocardial infarction is related to abnormalities in compliance, ventricular end-diastolic pressure, end-diastolic volume, and clinical manifestations of heart failure.
4. Exercise and atrial pacing can produce clinical and hemodynamic abnormalities.
5. The ejection fraction is significantly related to the slope of the ventricular function curve.
6. Angiographic abnormalities of left ventricular wall motion can be increased with atrial pacing and reduced with nitroglycerin or epinephrine.

Literature Cited

1. Rackley, C. E., Russell, R. O. Jr. 1973. *Coronary Care: Invasive Techniques For Hemodynamic Measurements.* New York: Am. Heart Assoc.
2. Russell, R. O. Jr., Rackley, C. E., Eds. 1974. Techniques and procedures. Chap. 2. *In Hemodynamic Monitoring in a Coronary Intensive Care Unit,* Mount Kisco, New York: Futura
3. Swan, H. J. C. et al. 1970. Catheterization of the heart in man with use of a flow-directed balloon-tipped catheter. *N. Engl. J. Med.* 283:447
4. Cohn, J. N., Khatri, I. M., Hamosh, P. 1970. Bedside catheterization of the left ventricle. *Am. J. Cardiol.* 25:66
5. Pombo, J. F., Russell, R. O. Jr., Rackley, C. E., Foster, G. L. 1971. Comparison of stroke volume and cardiac output determination by ultrasound and dye dilution in acute myocardial infarction. *Am. J. Cardiol.* 27:630
6. Ganz, W., Swan, H. J. C. 1972. Measurement of blood flow by thermodilution. *Am. J. Cardiol.* 29:241
7. Russell, R. O. Jr., Hunt, D., Potanin, C., Rackley, C. E. 1972. Hemodynamic monitoring in a coronary intensive care unit: Clinical applications. *Arch. Intern. Med.* 130:370
8. Russell, R. O. Jr. et al 1970. Effects of increasing left ventricular filling pressure in patients with acute myocardial infarction. *J. Clin. Invest.* 49:1539
9. Parmley, W. W., Diamond, G., Tomoda, H., Forrester, J. S., Swan, H. J. C. 1972. Clinical evaluation of left ventricular pressures in myocardial infarction. *Circulation* 45:358
10. Kiely, J., Kelly, D. T., Taylor, D. R., Pitt, B. 1973. The role of furosemide in the treatment of left ventricular dysfunction associated with acute myocardial infarction. *Circulation* 48:581
11. Hodges, M., Friesinger, G., Riggins, R., Dagenais, G. R. 1972. Effects of intravenously administered digoxin on mild left ventricular failure in acute myocardial infarction. *Am. J. Cardiol.* 29:749

12. Amsterdam, E. A. et al 1974. Hemodynamic effects of propranolol in acute myocardial infarction: Beneficial and deleterious actions. *Clin. Res.* 22:257A
13. Sones, F. M. Jr., Shirley, E. K. 1962. Cine coronary arteriography. *Mod. Conc. Cardiovasc. Dis.* 31:735
14. Judkins, M. P. 1968. Percutaneous transfemoral selective coronary arteriography. *Radiol. Clin. N. Am.* 3:467
15. Rackley, C. E., Hood, W. P. Jr., Rolett, E. L., Young, D. T. 1970. Left ventricular end-diastolic pressure in chronic heart disease. *Am. J. Med.* 48:310
16. Smith, M., Russell, R. O. Jr., Moraski, R. E., Feild, B. J., Rackley, C. E. 1974. Left ventricular A-wave amplitude in post-myocardial infarction patients. *Am. J. Cardiol.* 33:370
17. Mason, D. T. 1973. Regulation of cardiac performance in clinical heart disease: Interactions between contractile state mechanical abnormalities and ventricular compensatory mechanisms. *Am. J. Cardiol.* 32:437
18. Dodge, H. T., Sandler, H., Ballew, D. W., Lord, J. D. Jr. 1960. The use of biplane angiocardiography for the measurement of left ventricular volume in man. *Am. Heart J.* 60:762
19. Greene, D. G., Carlisle, R., Grant, C., Bunnell, I. L. 1967. Estimation of left ventricular volume by one-plane cineangiography. *Circulation* 35:61
20. Rackley, C. E., Behar, V. S., Whalen, R. E., McIntosh, H. D. 1967. Biplane cineangiographic determinations of left ventricular function: Pressure-volume relationships. *Am. Heart J.* 74:766
21. Rackley, C. E., Dodge, H. T., Coble, Y. D. Jr., Hay, R. E. 1964. A method for determining left ventricular mass in man. *Circulation* 29:666
22. Sandler, H., Dodge, H. T. 1963. Left ventricular tension and stress in man. *Circ. Res.* 13:91
23. Bristow, J. D., Van Zee, B. E., Judkins, M. P. 1970. Systolic and diastolic abnormalities of the left ventricle in coronary artery disease. *Circulation* 42:219
24. Smith, M., Russell, R. O. Jr., Feild, B. J., Rackley, C. E. 1974. Left ventricular specific compliance and abnormally contracting segments following myocardial infarction. *Chest* 65:368
25. Herman, M. V., Heinle, R. A., Klein, M. D., Gorlin, R. 1967. Localized disorders in myocardial contraction. *N. Engl. J. Med.* 277:222
26. Hamilton, G. W., Murray, J. A., Kennedy, J. W. 1972. Quantitative angiocardiography in ischemic heart disease. The spectrum of abnormal left ventricular function and role of abnormally contracting segments. *Circulation* 45:1065
27. Feild, B. J., Russell, R. O. Jr., Dowling, J. T., Rackley, C. E. 1972. Regional left ventricular performance in the year following myocardial infarction. *Circulation* 46:679
28. Messer, J. V., Levine, H. J., Wagman, R. J., Gorlin, R. 1963. Effect of exercise on cardiac performance in human subjects with coronary artery disease. *Circulation* 28:404
29. Cohen, L. S., Elliott, W. C., Rolett, E. L., Gorlin, R. 1965. Hemodynamic studies during angina pectoris. *Circulation* 31:409
30. Parker, J. O., diGiorgi, S., West, R. O. 1966. A hemodynamic study of acute coronary insufficiency precipitated by exercise with observations on the effects of nitroglycerin. *Am. J. Cardiol.* 17:470
31. Sowton, G. E., Balcon, R., Cross, D., Frick, M. G. 1967. Measurement of the angina threshold using atrial pacing: New technique for the study of angina pectoris. *Cardiovasc. Res.* 1:301
32. O'Brien, K, P., Higgs, L. M., Glancy, D. L., Epstein, S. E. 1969. Hemodynamic accompaniments of angina: A comparison during angina induced by exercise and by atrial pacing. *Circulation* 39:735
33. Parker, J. O., Ledwich, J. R., West, R. O., Case, R. B. 1969. Reversible cardiac failure during angina pectoris: Hemodynamic effects of atrial pacing in coronary heart disease. *Circulation* 39:745
34. Linhart, J. W., Hildner, F. J., Barold, S. S., Lister, J. W., Samet, P. 1969. Left heart hemodynamics during angina pectoris induced by atrial pacing. *Circulation* 40:483
35. Khaja, F., Sanghvi, V. R., Mark, A., Parker, J. O. 1971. Effects of volume expansion on anginal threshold. *Circulation* 43:824
36. Feild, B. J., Russell, R. O. Jr., Smith, M., Rackley, C. E. 1973. Left ventricular function curves and ejection fraction in post-myocardial infarction patients. *Am. J. Cardiol.* 31:130
37. Cohn, P. F. et al 1973. Effect of angiographic contrast medium on left ventricular function in coronary artery disease. *Am. J. Cardiol.* 32:21
38. Ross, J. Jr., Braunwald, E. 1964. Study of left ventricular function in man by

increasing resistance to ventricular ejection with angiotensin. *Circulation* 29:739

39. Linhart, J. W., Hildner, F. J., Barold, S. S., Samet, P. 1969. Myocardial function in patients with coronary artery disease. *Am. J. Cardiol.* 23:379
40. Dwyer, E. M. 1970. Left ventricular pressure-volume alterations and regional disorders of contraction during myocardial ischemia incuded by atrial pacing. *Circulation* 42:1111
41. Pasternac, A., Gorlin, R., Sonnenblick, E. H., Hoft, J. I., Kemp, H. G. 1972. Abnormalities of ventricular motion induced by atrial pacing in coronary artery disease. *Circulation* 45:1195
42. Pine, R. et al 1973. Detection of reversible ventricular contraction abnormalities with nitroglycerin: Correlation with post coronary bypass ventriculography. *Circulation* 48:IV–104
43. Dumesnil, J. G. et al 1973. Regional left ventricular wall dynamics before and after the administration of oral nitroglycerin. *Circulation* 48:IV–104
44. Horn, H. R., Teichholz, L. E., Cohn, P. F., Herman, M. V., Gorlin, R. 1974. Augmentation of left ventricular contraction pattern in coronary artery disease by inotropic catecholamine: The epinephrine ventriculogram. *Circulation* 49:1063
45. Freis, E. D., Schnapper, H. W., Johnson, R. L., Schreiner, G. E. 1952. Hemodynamic alterations in acute myocardial infarction: I. Cardiac output, mean atrial pressure, total peripheral resistance, central and total blood volumes, venous pressure and average circulation time. *J. Clin. Invest.* 31:131
46. Gilbert, R. P., Goldberg, M., Griffin, J. 1954. Circulatory changes in acute myocardial infarction. *Circulation* 9:847
47. Gammill, J. F., Applegarth, J. J., Reed, C. E., Fernald, J. D., Antenucci, A. J. 1955. Hemodynamic changes following acute myocardial infarction using the dye injection method for cardiac output determination. *Ann. Intern. Med.* 43:100
48. Gunnar, R. M. et al 1966. Myocardial infarction with shock. Hemodynamic studies and results of therapy. *Circulation* 33:753
49. Smith, H. J., Oriol, A., March, J., McGregor, M. 1967. Hemodynamic studies in cardiogenic shock: Treatment with isoproterenol and metaramine. *Circulation* 35:1084
50. MacKenzie, A. J. et al 1964. Circulatory and respiratory studies in myocardial infarction and cardiogenic shock. *Lancet* 2:825
51. Cohn, J. N., Luria, M. H., Daddario, R. C., Tristani, F. E. 1967. Studies in clinical shock and hypotension: V. Hemodynamic effects of dextran. *Circulation* 35:316
52. Nixon, P. G. F., Ikram, H., Morton, S. D. 1966. Infusion of dextrose solution in cardiogenic shock. *Lancet* 1:1077
53. Cohn, J. N., Ibrahim, M., Khartri, M. D., Tristani, F. E. 1968. Left and right ventricular filling pressures in clinical shock. *Ann. Intern. Med.* 68:1153
54. Hamosh, P., Cohn, J. N. 1971. Left ventricular function in acute myocardial infarction. *J. Clin. Invest.* 50:523
55. Rapaport, E., Scheinman, M. 1969. Rationale and limitations of hemodynamic measurements in patients with acute myocardial infarction. *Mod. Conc. Cardiovasc. Dis.* 38:55
56. Forrester, J. S., Diamond, G., McHugh, T. J., Swan, H. J. C. 1971. Filling pressure in the right and left side of the heart in acute myocardial infarction. A reappraisal of central venous pressure monitoring. *N. Engl. J. Med.* 285:190
57. Rackley, C. E., Russell, R. O. Jr. 1972. Left ventricular function in acute myocardial infarction and its clinical significance. *Circulation* 45:231
58. Cohn, J. N., Guiha, N. H., Broder, M. I., Limas, C. J. 1974. Right ventricular infarction: Clinical and hemodynamic features. *Am. J. Cardiol.* 33:209
59. Rackley, C. E., Russell, R. O. Jr. 1974. Right ventricular function in acute myocardial infarction. *Am. J. Cardiol.* 33:927
60. Rutherford, B. D., McCann, W. D., O'Donovan, T. P. B. 1971. The value of monitoring pulmonary artery pressure for early detection of left ventricular failure following myocardial infarction. *Circulation* 43:655
61. Kaltman, A. J., Herbert, W. H., Conroy, R. J., Kossman, C. E. 1966. Gradient in pressure across the pulmonary vascular bed during diastole. *Circulation* 34:377
62. Falicov, R. E., Resnekov, L. 1970. Relationship of the pulmonary end-diastolic pressure to the left ventricular end-diastolic and mean filling pressures in patients with and without left ventricular dysfunction. *Circulation* 42:65
63. Bouchard, R. J., Gault, J. H., Ross, J. Jr. 1970. Evaluation of pulmonary ar-

terial end-diastolic pressure as an estimate of left ventricular end-diastolic pressure in patients with acute and chronic alterations in left ventricular performance. *Am. J. Cardiol.* 25:86
64. Rahimtoola, S. et al 1972. Relationship of pulmonary artery to left ventricular diastolic pressures in acute myocardial infarction. *Circulation* 46:283
65. Russell, R. O. Jr., Hunt, D., Rackley, C. E. 1973. Left ventricular hemodynamics in anterior and inferior myocardial infarction. *Am. J. Cardiol.* 32:8
66. Scheinman, M., Evans, G. T., Weiss, A., Rapaport, E. 1973. Relationship between pulmonary artery end-diastolic pressure and left ventricular filling pressure in patients in shock. *Circulation* 47:317
67. Rackley, C. E., Russell, R. O. Jr., Mantle, J. A. 1974. Hemodynamic measurements in clinically uncomplicated myocardial infarction. See Ref. 2, Chap. 8
68. Rotman, M. et al 1974. Pulmonary arterial diastolic pressure in acute myocardial infarction. *Am. J. Cardiol.* 33:357
69. Scheidt, S., Ascheim, R., Killip, T. 1971. Shock after acute myocardial infarction. *Am. J. Cardiol.* 26:556
70. Ratshin, R. A., Rackley, C. E., Russell, R. O. Jr. 1972. Hemodynamic evaluation of left ventricular function in shock complicating myocardial infarction. *Circulation* 45:127
71. Rackley, C. E., Russell, R. O. Jr., Ratshin, R. A., Weber, K. T., Mantle, J. A. 1974. Cardiogenic shock in patients with myocardial infarction. See Ref. 2, Chap. 10
72. Rackley, C. E., Russell, R. O. Jr., Mantle, J. A. 1974. Clinical considerations of hemodynamic measurements and left ventricular function in myocardial infarction. See Ref. 2, Chap. 7
73. Greene, H. L., Kelly, D. T., Taylor, D. R., Pitt, B. 1974. Hemodynamic effects of plasma volume expansion and prognostic implications in acute myocardial infarction. *Circulation* 49:106
74. Rackley, C. E., Russel, R. O. Jr., Mantle, J. A. 1974. Hemodynamic measurements of heart failure in patients with myocardial infarction. See Ref. 2, Chap. 9
75. Karliner, J. S., Braunwald, E. 1972. Present status of digitalis treatment of acute myocardial infarction. *Circulation* 45:891
76. Chatterjee, K. et al 1973. Hemodynamic and metabolic responses to vasodilator therapy in acute myocardial infarction. *Circulation* 48:1183
77. Gold, H. K., Leinbach, R. C., Sanders, C. A. 1972. Use of sublingual nitroglycerin in congestive failure following acute myocardial infarction. *Circulation* 46:839
78. Shillingford, J. P., Thomas, M. 1967. Cardiovascular and plumonary changes in patients with myocardial infarction treated in an intensive care and research unit. *Am. J. Cardiol.* 20:484
79. Fluck, D. C. et al 1967. Right heart pressures in acute myocardial infarction. *Brit. Heart J.* 29:748
80. Weil, H. M., Schubin, H. 1968. Shock following acute myocardial infarction: Current understanding of hemodynamic mechanisms. *Progr. Cardiov. Dis.* 11:1
81. Smith, M., Ratshin, R. A., Harrell, F. E. Jr., Russell, R. O. Jr., Rackley, C. E. 1974. Early sequential changes in left ventricular dimensions and filling pressure in post-myocardial infarction patients. *Am. J. Cardiol.* 33:363
82. Rackley, C. E., Dear, H. D., Baxley, W. A., Jones, W. B., Dodge, H. T. 1970. Left ventricular chamber volume, mass and function in severe coronary artery disease. *Circulation* 41:605
83. Hamby, R. I., Supta, M. P., Young, M. W. 1973. Clinical and hemodynamic aspects of single vessel coronary artery disease. *Am. Heart J.* 85:458
84. Moraski, R. E., Russell, R. O. Jr., Smith, M., Rackley, C. E. 1975. The relationship between one, two and three vessel coronary artery disease, left ventricular function and myocardial infarction. *Am. J. Cardiol.* Vol. 35
85. Falsetti, H. L., Geraci, A. R., Bunnell, I. L., Greene, D. G., Grant, C. 1971. Function of left ventricle and extent of coronary lesions: Failure of correlation in cine-angiographic studies. *Chest* 49:610
86. Gould, K. L., Lipscomb, K., Hamilton, G. W., Kennedy, J. W. 1973. Left ventricular hypertrophy in coronary artery disease. A cardiomyopathy syndrome following myocardial infarction. *Am. J. Med.* 55:595
87. Diamond, G., Forrester, J. S. 1972. Effect of coronary artery disease and acute myocardial infarction on left ventricular compliance in man. *Circulation* 45:11
88. Porter, C. M., Baxley, W. A., Eddleman, E. E. Jr., Frimer, M., Rackley, C. E. 1971. Left ventricular dimensions and dynamics of filling in patients with

gallop heart sounds. *Am. J. Med.* 50:721

89. Rogers, W. J., Moraski, R. E., Russell, R. O. Jr., Rackley, C. E. 1974. Evaluation of left ventricular contractile state in coronary artery disease. *Clin. Res.* 22:12a
90. Herman, M. V., Gorlin, R. 1969. Implications of left ventricular asynergy. *Am. J. Cardiol.* 22:538
91. Foster, G. L., Reeves, T. J. 1964. Hemodynamic responses to exercise in clinically normal middle-aged men and in those with angina pectoris. *J. Clin. Invest.* 43:1758
92. Rackley, C. E. et al Catheterization evaluation of cardiac function in acute and chronic coronary artery disease. *Am. J. Cardiol.* In press

Copyright 1975. All rights reserved

DENTAL CARIES

❖7136

R. J. Gibbons, Ph.D. and J. van Houte, D.M.D., Ph.D.
Forsyth Dental Center, Boston, Massachusetts 02115

INTRODUCTION

Dental caries is one of the most prevalent infectious diseases of man, and for practical purposes it can be considered ubiquitous in civilized populations. Consequently the cost of its treatment ranks among the most expensive of bacterial infections; it also results in considerable loss of time and productivity (1). Dental caries may be defined as the localized destruction of tooth tissue by bacterial action. The cavities begin as demineralized foci on the peripheral surface of the enamel covering the crowns of teeth. Early lesions are recognized clinically as "white spots" and they eventually progress into the dentin and reach the pulp (2). The demineralization is thought to be caused by acids resulting from the bacterial fermentation of dietary carbohydrates, although other mechanisms of tooth destruction have been postulated (2–4). Dissolution of the hydroxyapatite crystals seems to precede the loss of the organic matrices of both enamel and dentin (2). Not all surfaces of a tooth are equally afflicted; rather, areas which are relatively protected from cleansing action such as occlusal fissures, pits, and areas below the contact points between teeth are more frequently attacked than the smooth, nonretentive buccal, lingual, or facial surfaces.

ETIOLOGY

Shortly after it was recognized that bacteria could cause diseases in man and animals, W. D. Miller (5) demonstrated that salivary bacteria produced acids in sufficient quantity from carbohydrate fermentation to decalcify teeth. This and other experiments led him to postulate that bacteria were the causative agents of dental caries. There are now more convincing data which substantiate the bacterial etiology. Orland and co-workers (6) demonstrated that germ-free rats remained free of caries even when maintained on a high carbohydrate diet, but they developed decay when infected with certain organisms (7). Antibiotics also can completely control this disease when administered to experimental animals (8–11), and these agents also affect human tooth decay (12). In addition, unerupted teeth do not develop decay; rather direct exposure to the oral environment is essential (2). More recently, Keyes (13) demonstrated that in populations of conventional hamsters and rats, this disease can be transmitted between animals. This observation underscores its infectious nature.

Epidemiologic studies in humans have shown that frequent ingestion of sucrose-containing foods increases dental caries activity (14, 15). Sucrose may a priori be considered important in dental caries since it is readily fermented and comprises 30–50% of the carbohydrate consumed in affluent societies; most of the remaining carbohydrate is derived from starches. The carbohydrate must be acted upon by the bacteria locally in the mouth, for when carbohydrate-containing diets are administered by stomach tube to experimental animals, dental caries does not develop (16). It is not clear whether other readily fermented sugars such as glucose and fructose would be as conducive to human tooth decay as sucrose since they are not major dietary constituents. However, caries will not only develop in animals fed diets with sucrose, but also with these monosaccharides or lactose (17–21). Starch is less conducive, presumably because it must be hydrolyzed before it can be fermented, and because it diffuses with difficulty into the bacterial masses on teeth. The ability of lactose to serve as a substrate for bacterial glycolysis may be of significance in "milk bottle" caries of infants.

DENTAL PLAQUE

Carious lesions develop under dense bacterial masses, termed dental plaques, which are adherent to the tooth surface. The bacteria are separated from the enamel substance by a thin amorphous organic film termed the acquired pellicle (22). The pellicle is believed to be formed by the selective adsorption of salivary components to the hydroxyapatite crystals of enamel. Salivary and bacterial products also comprise part of the interbacterial matrix which maintains the integrity of the accumulated bacterial masses. Many different bacterial species colonize teeth, and their distribution varies from person to person, from tooth to tooth, and even from one area to another on the same tooth (22, 23). Hence the term "dental plaque" only vaguely describes the nature of these bacterial accumulations.

Streptococci and gram-positive filamentous bacteria appear to be among the most prominent organisms on coronal tooth surfaces where they number over 10^{10}/gm wet weight of plaque (23). They subject localized areas of the tooth to high concentrations of bacterial end products. Measurements of the pH of dental plaques in situ by Stephan (24) have indicated that acidities as high as pH 4 can be attained within minutes following a glucose rinse. Such acidities are considered sufficient to cause demineralization of the enamel surface. Stephan (24) also observed that the pH of plaques differed between individuals and was related to their dental caries experience.

Many bacteria which colonize the teeth are known to synthesize intra- and extracellular polysaccharides from dietary carbohydrates (25). Certain of these, i.e. intracellular polysaccharides similar to glycogen, and extracellular levans, may be readily metabolized to organic acids in the absence of environmental substrates (25). They serve, therefore, as bacterial energy reserves. In view of the intermittent intake of dietary carbohydrates, acids produced from such storage polysaccharides prolong the duration of acid production by plaque bacteria. This may be especially relevant at night when the flow of buffering saliva is greatly diminished. Bacteria which synthesize glycogen also ferment exogenous carbohydrates at a higher rate in vitro, presumably because ATP generated from glycolysis is required for glycogen synthesis (26). The proportions of bacteria which synthesize large quantities of intracellular glycogen are increased on the teeth of individuals with high dental caries activity (27–29).

BACTERIAL SPECIFICITY IN EXPERIMENTAL DENTAL CARIES

Experiments demonstrating the infectious nature of dental caries have indicated that considerable bacterial specificity is involved in the initiation of this disease in rodents. Keyes (13) observed that a strain of albino hamsters did not develop caries when maintained on a caries-conducive high sucrose diet. However when these animals were caged together with caries-active golden hamsters, or were inoculated intra-orally with the fecal flora of these animals, they developed rampant decay. Transmissibility of dental caries has also been observed in strains of conventional rats (13, 30). Fitzgerald & Keyes (31) subsequently isolated specific streptococci from rodent carious lesions which initiated rampant decay when inoculated intraorally in "caries-inactive" albino hamsters. These caries-inducing streptococci are now considered strains of *Streptococcus mutans* (32, 33). Although the albino hamsters harbored a complex bacterial flora, they evidently lacked bacteria capable of initiating dental caries. Concurrently, Fitzgerald and co-workers demonstrated that various bacterial species introduced as monocontaminants in gnotobiotic rats exhibited marked differences in their cariogenic potential (34).

These studies provided animal models that could be utilized to assess the cariogenic potential of oral bacteria. The subsequent testing of a number of bacterial species in conventional and gnotobiotic animals has shown that a spectrum of cariogenic potential exists among bacteria (Table 1). In hamsters, whose teeth do not contain occlusal fissures, a limited number of bacterial species have been found capable of initiating decay. Strains of *S. mutans*, isolated from either humans or rodents, consistently initiate decay affecting smooth enamel surfaces provided the animals are maintained on a diet rich in sucrose. Various *Actinomyces* species are able to elicit cemental carious lesions and to induce the breakdown of the supporting structures of the teeth. In contrast, several other species, including acidogenic organisms, are ineffective in initiating decay in the hamster (Table 1).

Analogous observations have been made with rats (Table 1). The molar teeth of rats have occlusal fissures which are comparable to those of human molars, and they are highly prone to decay. Strains of *S. mutans* almost always produce decay in pits and fissures, as well as on smooth enamel surfaces in both gnotobiotic and conventional rats. This organism will also induce caries in monkeys and gerbils (49, 50). On the other hand, only a few of the tested strains of *S. sanguis*, *S. mitis*, enterococci, lactobacilli, and *Actinomyces* initiate enamel carious lesions, and these are almost always restricted to fissures (Table 1). The filamentous organisms, however, are adept at colonizing the gingival crevice area, and quite consistently induce cemental or root surface lesions. Thus the formation of carious lesions on different tooth surfaces may involve different bacterial species.

Animals infected with *S. mutans* or certain *Actinomyces* species develop significant bacterial accumulations on their teeth, whereas animals infected with organisms of little or no cariogenic potential generally do not (Table 1). Therefore, an organism must accumulate to a critical mass on the tooth surface to be pathogenic. This property per se is not the sole determinant of cariogenicity, since strains of *S. sanguis* and *Actinomyces* may accumulate without necessarily eliciting enamel or cemental decay (Table 1). In vitro experiments (51) and direct measurements of plaque pHs in experimental animals (52) have suggested that these organisms do not create the intense acid environment necessary for enamel demineralization. The ability of the *Actinomyces* to produce cemental caries may be because this tissue is less mineralized than enamel.

Studies with humans and experimental animals have shown that the caries experience of the offspring tends to mimic that of the parents (53–58). The results from these and other familial studies (53, 59) have been interpreted as suggesting that genetic factors significantly influenced caries susceptibility. It had been assumed that the populations studied were equally infected and challenged; or such differences were sought, but could not be demonstrated. However, it is now clear that rodent populations can be infected to different degrees by caries-inducing bacteria, and that caries activity can be transmitted from parents to their offspring (13, 30, 35, 60).

Table 1 Experimental dental caries initiated by various bacteria[a]

	"Pathogen-Free" Hamster				Germ-Free or Conventional Rat				
	Bacterial Accumulation		Smooth Enamel Lesions	Root Surface Lesions	Bacterial Accumulation		Smooth Enamel Lesions	Fissure Lesions	Root Surface Lesions
Organisms	Coronal	Gingival			Coronal	Gingival			
S. mutans	4+	1+	4+	1+	4+	1+	3+	4+	1+
Filamentous organisms including: *Actinomyces viscosus* *Actinomyces naeslundii* *Rothia* and *Nocardia* species	1+	4+	—	4+	1+	4+	—	– to 1+	4+
S. salivarius	?	?	– to 2+	– to 2+	±	– to 3+	—	—	– to 2+
S. sanguis	?	?	?		– to 2+		—	– to 1+	—
S. mitis	—	—	—	—	—	—	—	– to 1+	—
Enterococci	—	—	—	—	±	—	—	– to 1+	—
Various strains of lactobacilli	—	—	—	—	±	—	—	– to 2+	—
Various diphtheroids	—		—	—	—	—	—	—	—
Strains of:									
Neisseria	?	?	—	—	?	?	—	—	—
Fusiforms	?	?	—	—	?	?	—	—	—
gram-rods	?	?	—	—	?	?	—	—	—
and others	?	?	—	—	?	?	—	—	—

[a] Data compiled from references 33–48.

Humans have also been found to be infected with different levels of *S. mutans,* and this correlates with their caries experience. Consequently, the earlier familial studies do not necessarily indicate that genetic factors play a significant role. Other studies comparing the caries experiences of identical and nonidentical twins also suggest that genetic influences (61) do exist. However the influence of genetic factors in the population at large appears to be over shadowed by the severe environmental challenge.

EVIDENCE OF BACTERIAL SPECIFICITY IN HUMAN DENTAL CARIES

Lactobacilli

Ever since it was recognized that specific microbial agents were etiologic agents of various infectious diseases, it has been suspected that some bacteria were more important than others in initiating human dental decay. The reports by Kligler in 1915 (62, 63) that plaques on carious surfaces had higher proportions of lactobacilli than plaques on caries-free surfaces prompted investigators to study the relationship of lactobacilli with this disease. In comparisons of groups, it was found that the numbers of lactobacilli in saliva or on the tooth surface statistically correlated with caries activity but this positive relationship was often absent in individual subjects (64–67). The presence of lactobacilli in the mouth was also reported to precede clinically detectable carious lesions (66), and their numbers in saliva were positively influenced by the quantity of dietary carbohydrate ingested (68). However, more recent studies have shown that the bacterial populations of saliva are not representative of those on the tooth surface (69, 70), and that lactobacilli constitute only a very small fraction of the plaque flora (71–74). In addition, their numbers on teeth tend to increase only after decay has begun, and hence a carious lesion may be a prerequisite for their colonization. The observation that lactobacilli disappear from the mouth following restoration of lesions is consistent with this concept (75, 76). This relationship may be explained by the fact that lactobacilli have a relatively feeble ability to sorb to teeth, and their numbers in saliva available for sorption are generally low (73). Consequently, a stagnant, protected area such as a carious lesion seems to provide a sheltered niche which fosters their retention and subsequent growth. Although a few of the strains tested have induced decay in gnotobiotic rats (Table 1), the collective data available suggest that lactobacilli are not of major importance in the initiation of human carious lesions. On the other hand, their statistical association with caries activity, their presence within lesions themselves, and their recognized acidogenic and aciduric properties suggest that they may be important secondary invaders which contribute to the progression of tooth decay.

Streptococci

In the mid 1950s, Orland and co-workers observed that dental caries would develop in gnotobiotic rats in association with enterococci (7), thereby suggesting that these acidogenic and aciduric organisms might be important in the etiology of certain types of carious lesions. However, it has been shown that the incidence and numbers of enterococci in dental plaque and in other areas of the mouth are low (77–79), and that several strains lack demonstrable cariogenicity in gnotobiotic rats (O. G. Gold et al, personal communication). Enterococci have been recovered in significant proportions from human carious lesions, particularly from the advancing front (80). Like the lactobacilli, these organisms may be significant secondary invaders, but there are few data which suggest their importance in human caries initiation.

In the 1960s, studies with pathogen-free hamsters and gnotobiotic rats revealed the high cariogenic potential of *S. mutans* (Table 1). Consequently there has been considerable interest in seeking possible associations of this organism with human dental caries activity. Utilizing plaque samples pooled from several tooth surfaces, a number of investigators have reported increased proportions of *S. mutans* associated with high caries activity (81–84). Other studies have demonstrated a strong positive association between the proportion of this organism and human caries experience when plaque samples obtained from individual tooth surfaces were analyzed (85–87). The number of tooth surfaces infected with *S. mutans* also parallels dental caries experience (86, 88), and such infected surfaces have been found to frequently develop carious lesions within 18 months to 2 years, whereas surfaces on which *S. mutans* could not be detected generally remain sound (74).

S. mutans colonizes teeth in a highly localized manner. It may be more frequently isolated from areas of teeth which are recognized to be highly caries susceptible than from less susceptible areas (86–88). *S. mutans* has also been isolated from almost all human enamel carious lesions examined (89–91), and it has been reported to be present in a significant percentage of cemental lesions (92, 93). The carious lesion itself appears to constitute an important source of this organism for infecting other tooth surfaces, since sound teeth adjacent to carious teeth are more commonly infected than are distant ones (90). Restoration of carious lesions also reduces the percentages of neighboring tooth surfaces which are infected with this organism (94). Collectively, the data indicate that *S. mutans* must be considered an important organism in the initiation of carious lesions on enamel surfaces, but unequivocal evidence concerning its direct involvement in human disease has yet to be obtained.

Although *S. sanguis* and *S. mitis* are prominent colonizers of teeth, and a few strains have exhibited cariogenic potential in rodents (33), their proportions have been found to correlate inversely with caries activity in both humans (85, 95) and monkeys (96). Their pattern of colonization on given areas of the tooth also does not parallel the recognized caries susceptibility of these areas. Consequently, these streptococci seem to be much less important than *S. mutans* in human decay. Strains of *S. salivarius* also possess cariogenic activity when tested in gnotobiotic rats (41, 97) and conventional hamsters (36). However, this organism constitutes only a small percentage of the bacterial flora of dental plaque (69, 70), and it is rarely isolated from carious lesions (91). Thus it is not likely to be an important etiologic agent.

It is of interest to note that *S. mutans, S. sanguis,* and *S. mitis* comprise a high percentage of the organisms associated with cases of human subacute bacterial endocarditis (98). All three species are prevalent on teeth, but *S. mutans* and *S. sanguis* are present in only low numbers in the intestinal tract (99, 100). Their distribution, and the association of *S. mutans* with human carious lesions, suggest that infected teeth are a major portal of entry of these organisms into the blood stream.

Filamentous Bacteria

Several types of filamentous bacteria have been observed to initiate root surface carious lesions as well as the destruction of the supporting tooth structures in hamsters and rats (Table 1). Some strains may also induce decay in the molar fissures of gnotobiotic rats (47). *Actinomyces viscosus, A. naeslundii,* and *Rothia* and *Nocardia* spp. have been studied in this regard. These organisms are found in human plaque and dental calculus in significant numbers (23, 101), and they have been isolated in high proportions from advanced human dentinal lesions (91). They appear to be especially prominent within decayed cementum overlying the roots of teeth (44, 92, 93). Their demonstrable pathogenic potential in experimental animals, and their numerical prominence suggest that they are important pathogens for the initiation of carious lesions on cemental surfaces and for the induction of certain types of periodontal pathology. Thus far definitive efforts to study the association of filamentous organisms with these infectious processes in humans have been hampered by the lack of practical techniques for their enumeration, and by difficulties encountered with their identification.

ECOLOGY OF STREPTOCOCCUS MUTANS

Studies in experimental animals (102) and in man (103) have shown that dietary sucrose favors the implantation of *S. mutans* in the mouth, and its accumulation on teeth. These effects are due to the organism's ability to synthesize extracellular glucans and fructans specifically from sucrose (22, 25, 104), which may account in part for the pronounced caries-promoting potential of this sugar. The glucans formed consist of polymers which contain mainly α-1,6 linkages similar to classical dextrans, and of polymers which are predominantly α-1,3 linked. The latter polysaccharides, which are less commonly produced by bacteria, have been referred to as "mutan." The fructans produced are similar to levan. An intact sucrose molecule is required for the synthesis of both the extracellular glucans and fructans because the enzymes involved are glucosyl- and fructosyl-transferases which utilize the energy contained in the disaccharide bond of sucrose for polymerization.

The glucans appear to be more important in plaque formation by *S. mutans* than are the fructans since enzymes which hydrolyze either dextran or mutan are effective in inhibiting these streptococci from accumulating on the teeth of experimental animals, and from forming adherent deposits in vitro (105, 106). Low molecular weight dextrans, which alter glucan polymerization, exert a similar effect (107). The glucans, and particularly mutan, tend to be less soluble than the fructans, and they are degraded comparatively slowly by oral bacteria. Consequently the glucans appear to persist for longer periods of time in mixed human plaque, thereby more effectively serving as matrix components to maintain the integrity of *S. mutans* microcolonies (22).

Glucose-grown cells of *S. mutans* possess receptor sites which specifically bind dextran molecules, and this results in aggregation (108). Other types of plaque bacteria, including strains of *S. sanguis* which also synthesize extracellular glucans, do not display this characteristic; evidently, they do not possess the surface receptors required for dextran binding (108). This is consistent with the observation that their colonization of teeth is not dependent upon dietary sucrose (22).

Glucan synthesis fosters the attachment of *S. mutans* to surfaces, which explains the enhancing effect of sucrose on its colonization (22, 109–111). The attachment of most *S. mutans* cells to hydroxyapatite surfaces appears to involve an initial phase of reversible sorption. De novo synthesis of polymers on the cell surface while the organisms are reversibly attached promotes more permanent binding (110, 111). *S. mutans* appears to require a nondesquamating surface to become established in the mouth. Studies by Carlsson have shown that it does not colonize the mouth of predentate infants; rather, infants become infected shortly after the teeth have erupted (112). *S. mutans* can also colonize acrylic obturators placed into the mouths of infants with cleft palates prior to tooth eruption (R. J. Berkowitz and H. V. Jordan, personal communication). To date, *S. mutans* has only been isolated from man and certain animals. Consequently infants most likely become infected from their parents, or from other individuals with which they have frequent contact. One may suspect that infants become more easily infected than older children and adults. For example, Jordan and co-workers (113) reported that streptomycin-resistant strains of *S. mutans* implanted into the mouths of mothers were not readily transmitted to their spouses or to older children. *S. mutans* also does not appear to spread readily from one tooth surface to another within the mouths of older children, adults, or in experimental animals. This is indicated by the highly localized pattern of colonization of the organism on teeth (86–83), and by the lack of transmission of artificially implanted streptomycin-resistant strains from one side of the mouth to the other (114). The reasons which account for the low rates of transmission of *S. mutans* between individuals and between teeth include: (*a*) its feeble capacity to adsorb to teeth and other oral surfaces resulting in its rapid clearance (115–117), and (*b*) its low salivary concentrations. It has been shown, for example, that approximately 10^5 cells of *S. mutans* per ml of saliva are required to have 1 cell sorb to a smooth tooth surface (115). In most individuals, the salivary concentrations of *S. mutans* are below this level. In addition, as previously discussed, most cells sorb reversibly to teeth and their firm attachment requires sucrose derived from the diet.

S. mutans is highly acidogenic and aciduric. Most strains synthesize significant quantities of intracellular glycogen, which further contributes to their acidogenic potential (118). Strains considered *S. mutans* are heterogeneous in a number of characteristics. Four groups have been recognized on the basis of guanine-cytosine DNA base ratios (119) which correspond to sero-groups elucidated by Bratthall (120). Furthermore, the groups differ in certain biochemical characteristics (121), and in isoenzyme and membrane protein patterns (122, 123). On the other hand, organisms of all groups have many phenotypic and ecologic similarities. They are all cariogenic when tested in experimental animals, they all synthesize extracellular glucans from sucrose, and with few exceptions, they ferment both mannitol and sorbitol. Organisms of all biotypes require teeth for colonization, and their presence is fostered by sucrose.

DENTAL CARIES PREVENTION

A surprisingly large segment of the public views dental caries as inevitable; yet it is clear that this disease can be prevented. Since its development requires the interactions of specific types of bacteria, dietary constituents, and a susceptible tooth surface, it is possible to interfere with the caries process by altering any of these parameters. A combination of restriction of the frequency and quantity of dietary carbohydrates ingested, particularly sucrose, the diligent adherence to oral hygiene procedures which remove bacterial accumulations from the teeth, and the administration of fluoride, particularly during the time of tooth development, will virtually completely prevent the disease (4, 124). Unfortunately, it has been found impossible to implement these measures as public health procedures. It is difficult to alter general eating habits of the population, and to achieve the proper motivation and knowledge required to mechanically control the bacterial deposits on teeth. In addition, there has frequently been public reluctance to adopt procedures such as water fluoridation because of emotional arguments, and a lack of appreciation of the infectious and transmissible nature of the disease. In addition, dental caries is rarely life-threatening. Therefore research is being aimed at developing preventative procedures which require minimal cooperation and effort on the part of the public. The increasing costs of dental treatment, however, may eventually provide a stimulus for better acceptance of existing preventive measures.

APPROACHES FOR CONTROLLING DENTAL CARIES

Occlusal Sealants

It is well established that pits and fissures on the occlusal surfaces of teeth are among the most frequently attacked. Earlier investigators attempted to alter the morphology of occlusal molar surfaces by grinding, or by sealing narrow fissures with filling materials so as to make the areas more self cleansing. Recently, self-polymerizing and UV light polymerizing plastic sealing materials have become available which firmly adhere to tooth material (125–129). Their application to occlusal fissures is generally preceded by cleaning and etching of the enamel with acid. In some studies, sealants have been retained for periods of over a year, and they have completely protected the underlying surfaces against decay. For optimal protection, the surfaces should be sealed shortly after tooth eruption because they usually decay within the first few years. Sealants of various types are gaining greater acceptance in certain preventive dentistry programs, but problems of retention still exist. Sealants are unavailable to protect the interproximal and buccal-lingual surfaces of teeth. The application of these procedures on a public health scale would create severe manpower problems under the present dental care delivery system.

Dietary Alteration

It is possible to interfere with the caries process by modifying the diet. Several studies have shown that dietary carbohydrates, particularly sucrose, markedly increase dental caries activity (14, 15, 35, 124, 130, 131). Studies in both man and experimental animals have indicated that not only is the type and amount of carbohydrate ingested of importance, but also the frequency of its ingestion (14, 15, 35, 132). Consequently, if one has a "sweet tooth," it is better to satisfy the craving in one episode, rather than nibble at sweets over a prolonged period. Invert sugar, sorbitol, hydrogenated starches, and artificial sweeteners have been suggested as sugar alternatives (35, 124), but they have not gained wide public acceptance. Consideration has also been given to additives which might counteract the detrimental effects of sucrose. For example, addition of low molecular weight dextrans which interfere with glucan synthesis to sucrose-containing diets reduces plaque formation and dental decay in experimental animals infected with *S. mutans* (35, 107, 124), but the quantities required are impractically high. The incorporation of various inorganic and organic phosphates in sucrose-rich diets has also been found to reduce dental caries in experimental animals (4, 35, 124). Generally, the phosphorus content of the diet must be doubled to get a significant effect. The mode of action of phosphate supplements is unclear, but they act locally within the oral cavity, rather than systemically. The few clinical trials carried out in humans in which certain components of the diet have been supplemented with phosphates have not been encouraging.

Fluoride and Other Trace Elements

Of all the preventive measures available today, the ingestion of fluoridated water or fluoride supplements during the time of tooth development produces the greatest caries-reducing effects (4, 124, 133, 134). Although the mechanism of the protective action of fluoride is still not firmly established, it is clear that the presence of one part of fluoride per million in drinking water generally results in a reduction of the caries rate in permanent teeth of approximately 50%. The effect of water fluoridation on deciduous teeth is somewhat less. The protection is most pronounced on smooth, nonretentive tooth surfaces and is the least in occlusal fissures. Despite its beneficial effects, only about half of the US population presently consumes water which contains optimal levels of fluoride. Daily intake of fluoride in liquid or tablet form from birth appears to be somewhat more effective than water fluoridation, presumably because infants consume less water (135). Topically applied fluoride preparations are generally less effective, while fluoride-containing dentifrices are only marginally efficacious.

There are data which suggest that elements other than fluoride may influence dental caries (136, 137). This information stems from epidemiological surveys similar to those which led to the discovery of the effects of fluoride, and from animal studies. Elements such as strontium, molybdenum, boron, lithium, and vanadium have been studied in this regard, but knowledge of their effects is meager compared to the well-established beneficial effect of fluoride.

Attempts to Control Caries-Inducing Bacteria

Since dental caries is an infectious disease, considerable effort is being aimed at interfering with the bacteria involved. Purely mechanical removal of the bacterial deposits on teeth by procedures such as tooth brushing, use of dental floss, and interdental stimulators may be effective when methodically performed by highly motivated individuals. However, even in such individuals, effective removal of the bacterial deposits below the contact points between teeth and in occlusal fissures is difficult to attain. Thus while these procedures may be beneficial, they are ineffective as public health measures because of the effort, cost, and motivation required of the individual. The prevalence of dental caries in civilized populations is tangible evidence of their inadequacy.

Rather than mechanically removing bacteria, attempts have been made to inhibit their ability to produce acids by, for example, application of inhibitors of glycolytic enzymes (138). So far, all effective agents have proven too toxic for continuous intra-oral use because they also inhibit mammalian cell glycolysis.

As a consequence of the recent appreciation of the bacterial specificity involved in caries etiology, and particularly because of the apparent importance of *S. mutans*, efforts have been directed towards specifically controlling this organism. Knowledge of the role of extracellular glucans in the infection and accumulation of *S. mutans* on teeth has led to attempts to interfere with glucan synthesis. Glucan hydrolases (i.e. dextranases and "mutanase") prepared from various fungi, and administered via the diet or drinking water to experimental animals are effective in preventing the accumulation of *S. mutans* on teeth, as well as the decay initiated by this organism (105, 106). However, studies of the use of dextranase mouthwashes to remove dental plaques in humans have not been encouraging (4), and a clinical trial designed to evaluate the possible caries-reducing effects of dextranase has had limited success. Other enzymes have also been considered for possible use in controlling the bacterial deposits which form on teeth, but thus far they have not been shown to be efficacious. If enzymes are to be of practical use, difficult obstacles to overcome include: their mode of administration, their short contact time in the mouth, their slow diffusion into the plaque, their specificities, and the avoidance of continuous use.

Chemotherapeutic approaches for controlling *S. mutans* infections offer some promise (124, 139). It is well documented that when antibiotics directed against gram-positive bacteria are added to the diet of experimental animals, dental caries is almost completely prevented (8–11, 124). Humans who have received long term antibiotic therapy for various medical reasons also exhibit a significantly reduced caries experience (140). Experimental antibiotic-containing dentifrices have been reported to produce similar effects (12). It is obvious that long term use of antibiotics and other chemotherapeutic agents which are of general medical importance is inadvisible for controlling dental caries. Nevertheless, the possibility exists that short term administration of these agents may be of practical value, particularly in severely infected or handicapped individuals. For example, pregnant hamsters treated with antibiotics for a brief time period develop minimal caries activity, and their offspring also remain caries-free even when fed caries-conducive diets (13).

Evidently, short term chemotherapy leads to the elimination of a necessary bacterial component, for when the offspring of treated hamsters are caged with carious animals, they subsequently develop decay. It has also been reported that antibiotic treatment can reduce the populations of *S. mutans* to undetectable levels in hamsters (141). Thus short term antibiotic administration seems to have the potential for eliminating important cariogenic bacteria from the mouth. Antibiotics, e.g. vancomycin and kanamycin, are currently being studied for their effectiveness in controlling *S. mutans* infections in humans (142, 143; W. J. Loesche, personal communication). Their intense topical application to the teeth of children for periods up to a few weeks causes a reduction in the *S. mutans* populations which persists for several weeks or months after the treatment is discontinued. Other significant alterations in the oral flora have not been observed. The persistent effect of these agents on the *S. mutans* populations of teeth seems to be attributable to the organism's slow rate of interdental and inter-individual transmission, and because it has no major oral reservoir other than the teeth. Therefore, if it is eliminated from an infected tooth surface by a chemotherapeutic agent, the surface tends to remain free of the organism for appreciable periods of time. Because of the organism's ecology, simple disinfecting agents also offer promise. For example, a single topical application of Lugol's iodine solution eliminated this organism from about half of the infected tooth surfaces treated for a period of over 3 months (88). Some fluoride preparations applied to teeth have been reported to exert similar effects (144, 145).

Antimicrobial agents, including kanamycin and chlorhexidine derivatives, have been studied for their ability to inhibit the overall bacterial colonization of teeth when administered in mouth rinses, or in dentifrices (146–149). Chlorhexidine adsorbs to teeth, and is slowly released, which prolongs its effect. Daily use of chlorhexidine has been observed to markedly reduce the bacterial accumulations on teeth for long periods of time. Use of this agent may also lead to the elimination of *S. mutans*, but to date, it has not been clearly demonstrated that procedures which eliminate *S. mutans* will also cause a reduction in human dental decay.

It has been possible to immunize experimental animals with vaccines of caries-inducing bacteria, and to obtain a partial reduction in dental caries development under certain experimental conditions. Wagner immunized gnotobiotic rats with a vaccine of *S. faecalis* cells and observed a reduction in fissure carious lesions following mono-infection with this organism (62). Subsequently, several investigators obtained evidence which suggests that immunization of rats, hamsters, and monkeys with vaccines of *S. mutans* cells can reduce caries induced by this organism (150–153). Vaccines consisting of crude glucosyltransferase preparations derived from *S. mutans* have also been reported to be effective (154). In vitro experiments have shown that IgG and IgA antibodies to glucosyltransferases inhibit glucan synthesis, and antiserum added to sucrose cultures inhibits the formation of adherent deposits by *S. mutans* (109, 154–157). Some investigators have not observed positive effects of immunization, or have obtained variable results (152, 158). This may be due to failure to stimulate salivary or other antibodies of the proper protective type in the oral cavity, too great an infective challenge imposed, or other variables between experiments. Immunization studies offer much promise, but they are still in their infancy, and the immunization of humans is unlikely to be attempted within the next few years.

Literature Cited

1. Cooper, B. S., Worthington, N. L. 1973. *Soc. Secur. Bull. DHEW Pub. (SSA) 73–11700,* pp. 1–20
2. Burnett, G. W., Scherp, H. W. 1968. *Oral Microbiology and Infectious Disease,* 328–408. Baltimore: Williams and Wilkins
3. Mäkinen, K. K. 1972. *Int. Dent. J.* 22:363–86
4. Scherp, H. W. 1971. *Science* 173:1199–1205
5. Miller, W. D. 1890. *The Micro-Organisms of the Human Mouth.* Philadephia: S. S. White Mfg.
6. Orland, F. J. et al 1954. *J. Dent. Res.* 33:147–74
7. Orland, F. J. et al 1955. *J. Am. Dent. Assoc.* 50:259–72
8. McClure, F. J., Hewitt, W. L. 1946. *J. Dent. Res.* 25:441–43
9. Stephan, R. M. et al 1952. *J. Dent. Res.* 31:421–27
10. Rosen, S., Ragheb, H. S., Hoppert, C. A., Hunt, H. R. 1956. *J. Dent. Res.* 35:399–403
11. Shaw, J. H., Sweeney, E. A. 1957. *J. Dent. Res.* 36:349–60
12. Zander, H. A. 1950. *J. Am. Dent. Assoc.* 40:569–74
13. Keyes, P. H. 1960. *Arch. Oral Biol.* 1:304–20
14. Gustafsson, B. E. et al 1954. *Acta Odontol. Scand.* 11:232–388
15. Takeuchi, M. 1960. *Bull. Tokyo Dent. Coll.* 1:58–70
16. Kite, O. V., Shaw, J. H., Sognnaes, R. F. 1950. *J. Nutr.* 42:89–103
17. Gustafson, G., Stelling, E. M., Abramson, E., Brunius, E. 1955. *Odontol. Tidskr.* 63:506–23
18. Grenby, T. H. 1963. *Arch. Oral Biol.* 8:27–30
19. Shaw, J. H., Krumins, I., Gibbons, R. J. 1967. *Arch. Oral Biol.* 12:755–68
20. Shafer, W. G. 1949. *Science* 110:143–44
21. Frostell, G., Keyes, P. H., Larson, R. H. 1967. *J. Nutr.* 93:65–76
22. Gibbons, R. J., van Houte, J. 1973. *J. Periodontol.* 44:347–60
23. Socransky, S. S., Manganiello, A. D. 1971. *J. Periodontol.* 42:485–94
24. Stephan, R. M. 1944. *J. Dent. Res.* 23:257–66
25. Gibbons, R. J. 1968. *Caries Res.* 2: 164–71
26. van Houte, J., Winkler, K. C., Jansen, H. M. 1969. *Arch. Oral Biol.* 14:45–61
27. Gibbons, R. J., Socransky, S. S. 1962. *Arch. Oral Biol.* 7:73–80
28. Loesche, W. J., Henry, C. A. 1967. *Arch. Oral Biol.* 12:189–94
29. van Houte, J., Backer Dirks, O., de Stoppelaar, J. D., Jansen, H. M. 1969. *Caries Res.* 3:178–89
30. Larson, R. H. 1968. In: *Art and Science of Dental Caries Research,* ed. R. S. Harris, 257–75. New York: Academic
31. Fitzgerald, R. J., Keyes, P. H. 1960. *J. Am. Dent. Assoc.* 61:9–19
32. Carlsson, J. 1967. *Odontol. Revy* 18: 55–74
33. Guggenheim, B. 1968. *Caries Res.* 2: 147–63
34. Fitzgerald, R. J., Jordan, H. V., Stanley, H. R. 1960. *J. Dent. Res.* 39:923–35
35. Keyes, P. H. 1968. *J. Am. Dent. Assoc.* 76:1357–73
36. Krasse, B., Carlsson, J. 1970. *Arch. Oral Biol.* 15:25–32
37. Jordan, H. V., Keyes, P. H. 1965. *Am. J. Pathol.* 46:843–57
38. Socransky, S. S. 1970. *J. Dent. Res.* 49:203–22
39. Fitzgerald, R. J. 1968. *Caries Res.* 2: 139–46
40. Gibbons, R. J., Berman, K. S., Knoettner, P., Kapsimalis, B. 1966. *Arch. Oral Biol.* 11:549–60
41. Gibbons, R. J., Banghart, S. 1968. *Arch. Oral Biol.* 13:297–308
42. Wagner, M. 1966. *A study of the effects of specific immunization on experimental dental caries in the gnotobiotic rat.* PhD thesis. Purdue University
43. Rosen, S., Lenney, W. S., O'Malley, J. E. 1968. *J. Dent. Res.* 47:358–63
44. Jordan, H. V., Hammond, B. F. 1972. *Arch. Oral Biol.* 17:1333–42
45. Jordan, H. V., Keyes, P. H. 1966. Arch. Oral Biol. 11:793–801
46. Keyes, P. H. 1962. *Int. Dent. J.* 12: 443–64
47. Frank, R. M., Guillo, B., Llory, H. 1972. *Arch. Oral Biol.* 17:1249–53
48. Llory, H., Guillo, B., Frank, R. M. 1971. *Helv. Odontol. Acta* 15:134–38

49. Bowen, W. H. 1969. *Caries Res.* 3:227–37
50. Fitzgerald, D. B., Fitzgerald, R. J. 1965. *Arch. Oral Biol.* 11:139–40
51. Larje, O., Frostell, G. 1968. *Acta Pathol. Microbiol. Scand.* 72:463
52. Charlton, G., Fitzgerald, D. B., Keyes, P. H. 1971. *Arch. Oral Biol.* 16:655–61
53. Mandel, I. D., Zengo, A. N. 1973. In *Comparative Immunology of the Oral Cavity* ed. S. E. Mergenhagen, H. W. Scherp, 118–37. *DHEW Publication (NIH) 73–438*
54. Finn, S. B. 1965. *Caries-Resistant Teeth.* ed. G.E.W. Wolstenholme, M. O'Connor, 41–60. Boston: Little, Brown
55. Klein, H. 1946. *J. Am. Dent. Assoc.* 33:735–43
56. Rosen, S., Hunt, H. R., Hoppert, C. A. 1961. *Arch. Oral Biol.* 5:92–97
57. Hunt, H. R., Hoppert, C. A., Erwin, W. G. 1944. *J. Dent. Res.* 23:385–401
58. Shaw, J. H., Griffiths, D. 1960. *Arch. Oral Biol.* 3:15–27
59. Böök, J. A., Grahnén, H. 1953. *Odontol. Revy* 4:1–53
60. Larson, R. H., Simms, M. E. 1965. *Arch. Oral Biol.* 10:663–68
61. Horowitz, S. L., Osborne, R. H., DeGeorge, F. V. 1958. *Science* 128:300–1
62. Kligler, I. J. 1915. *J. Allied Dent. Soc.* 10:141–66
63. Kligler, I. J. 1951. *J. Allied Dent. Soc.* 10:282–330
64. Hadley, F. P. 1933. *J. Dent. Res.* 13:415–28
65. Snyder, M. L. 1942. *J. Am. Dent. Assoc.* 29:2001–11
66. Enright, J. J., Friesell, H. E., Trescher, M. S. 1932. *J. Dent. Res.* 12:759–827
67. Hemmens, E. S., Blayney, J. R., Bradel, S. F., Harrison, R. W. 1946. *J. Dent. Res.* 25:195–205
68. Jay, P., Hadley, F. P., Bunting, R. W., Koehne, M. 1936. *J. Am. Dent. Assoc.* 23:846–51
69. Krasse, B. 1954. *Odontol. Revy* 5:203–11
70. Gibbons, R. J., Kapsimalis, B., Socransky, S. S. 1964. *Arch. Oral Biol.* 9:101–3
71. Strålfors, A. 1950. *Odontol. Tidskr.* 58:155–341
72. Krasse, B. 1954. *Odontol. Revy* 5:241–61
73. van Houte, J., Gibbons, R. J., Pulkkinen, A. J. 1972. *Infec. Immunity* 6:723–29
74. Ikeda, T., Sandham, H. J., Bradley, E. L. Jr. 1973. *Arch. Oral Biol.* 18:555–66
75. Kesel, R. G., Shklair, I. L., Green, G. H., Englander, H. R. 1958. *J. Dent. Res.* 37:50–51
76. Shklair, I. L., Englander, H. R., Stein, C. M., Kesel, R. G. 1956. *J. Am. Dent. Assoc.* 53:155–58
77. Williams, N. B., Forbes, M. A., Blau, E., Eickenberg, C. F. 1950. *J. Dent. Res.* 29:563–70
78. Bahn, A. N., Shklair, I. L., Mazzarella, M., Calandra, J. C. 1960. *J. Dent. Res.* 39:686–87
79. Gold, O. G., Jordan, H. V., van Houte, J. 1974. *52nd Session of Int. Assoc. Dent. Res.* 195 (Abstr.)
80. Burnett, G. W., Scherp, H. W. 1951. *J. Dent. Res.* 30:766–77.
81. Krasse, B. et al 1968. *Arch. Oral Biol.* 13:911–18
82. Englander, H. R., Jordan, H. V. 1972. *J. Dent. Res.* 51:1505
83. Schamschula, R. G., Charlton, G. 1971. *Aust. Dent. J.* 16:77–82
84. Woods, R. 1971. *Aust. Dent. J.* 16:116–21
85. de Stoppelaar, J. D., van Houte, J., Backer Dirks, O. 1969. *Caries Res.* 3:190–99
86. Shklair, I. L., Keene, H. J., Cullen, P. 1974. *Arch. Oral Biol.* 19:199–202
87. Ikeda, T., Sandham, H. J. 1971. *Arch. Oral Biol.* 16:1237–40
88. Gibbons, R. J., DePaola, P. F., Spinell, D. M., Skobe, Z. 1974. *Infec. Immunity* 9:481–88
89. Littleton, N. W., Kakehashi, S., Fitzgerald, R. J. 1970. *Arch. Oral Biol.* 15:461–63
90. Shklair, I. L., Keene, H. J., Simonson, L. G. 1972. *J. Dent. Res.* 51:882
91. Loesche, W. J., Syed, S. A. 1973. *Caries Res.* 7:201–16
92. Sumney, D. L., Jordan, H. V. 1974. *J. Dent. Res.* 53:343–51
93. Syed, S. A., Loesche, W. J., Pape, H. R. 1974. *52nd Session of Int. Assoc. Dent. Res.* p. 71 (Abstr.)
94. Keene, H. J., Shklair, I. L., Hoerman, K. C. 1973. In *Comparative Immunology of the Oral Cavity* ed. S. E. Mergenhagen, H. W. Scherp, 71–117. *DHEW Publication (NIH) 73–438*
95. Loesche, W. J., Walenga, A., Loos, P. 1973. *Arch. Oral Biol.* 18:571–75
96. Bowen, W. H. 1965. *Int. Dent. J.* 15:12–53
97. Kelstrup, J., Gibbons, R. J. 1970. *Caries Res.* 4:360–77

98. deMoor, C. E., van Houte, J., de Stoppelaar, J. D. 1968. *Versl. Meded. Volksgezondh.* 20:1–11
99. Kilian, M., Theilade, E., Schiott, C. R. 1971. *Arch. Oral Biol.* 16:553–54
100. van Houte, J., Jordan, H. V., Bellack, S. 1971. *Infect. and Immun.* 4:658–59
101. Collins, P. A., Gerencser, M. A., Slack, J. M. 1973. *Arch. Oral Biol.* 18:145–53
102. Krasse, B. 1965. *Arch. Oral Biol.* 10:215–21
103. Krasse, B., Edwardsson, S., Svensson, I., Trell, L. 1967. *Arch. Oral Biol.* 12:231–36
104. Guggenheim, B. 1970. *Int. Dent. J.* 20:657–78
105. Fitzgerald, R. J., Keyes, P. H., Stoudt, T. H., Spinell, D. M. 1968. *J. Am. Dent. Assoc.* 76:301–4
106. Guggenheim, B., Regolati, B., Mühlemann, H. R. 1972. *Caries Res.* 6:289–97
107. Gibbons, R. J., Keyes, P. H. 1969. *Arch. Oral Biol.* 14:721–24
108. Gibbons, R. J., Fitzgerald, R. J. 1969. *J. Bacteriol.* 98:341–45
109. Mukasa, H., Slade, H. D. 1973. *Infec. Immunity* 8:555–62
110. Clark, W., Gibbons, R. J., Skobe, Z. 1974. *52nd Session of Int. Assoc. Dent. Res.* p. 242 (Abstr.)
111. van Houte, J., Upeslacis, V. N. 1974. *52nd Session of Int. Assoc. Dent. Res.* p. 242 (Abstr.)
112. Carlsson, J., Grahnén, H., Jonsson, G., Wikner, S. 1970. *Arch. Oral Biol.* 15: 1143–48
113. Jordan, H. V., Englander, H. R., Engler, W. O., Kulczyk, S. 1972. *J. Dent. Res.* 51:515–18
114. Edman, D. C., Keene, H. J., Shklair, I. L., Hoerman, K. C. 1972. *50th Session of Int. Assoc. Dent. Res.* p. 67 (Abstr.)
115. van Houte, J., Green, D. B. 1974. *Infec. Immunity* 9:624–30
116. Gibbons, R. J., van Houte, J. 1971. *Infec. Immunity* 3:567–73
117. van Houte, J., Gibbons, R. J., Pulkkinen, A. J. 1971. *Arch. Oral Biol.* 16:1131–41
118. van Houte, J., deMoor, C. E., Jansen, H. M. 1970. *Arch. Oral Biol.* 15:263–66
119. Coykendall, A. L. 1973. *Streptococcus mutans and Dental Caries, 3–6. DHEW Publication (NIH) 74–286*
120. Bratthall, D. 1970. *Odontol. Revy* 21: 143–52
121. Shklair, I. L., Keene, H. J. 1974. *Arch. Oral Biol.* In press
122. Brown, A. T., Patterson, C. E. 1972. *Infec. Immunity* 6:422–24
123. Hamada, S., Mizuno, J. 1974. *J. Dent. Res.* 53:547–53
124. Keyes, P. H. 1969. *J. Am. Dent. Assoc.* 79:1395–1404
125. Cueto, E. I., Buonocore, M. G. 1967. *J. Am. Dent. Assoc.* 75:121–28
126. Ripa, L. W., Cole, W. W. 1970. *J. Dent. Res.* 49:171–73
127. Silverstone, L. M. 1974. *Caries Res.* 8:2–26
128. Buonocore, M. G. 1970. *J. Am. Dent. Assoc.* 80:324–28
129. Reports of Councils and Bureaus. 1971. *J. Am. Dent. Assoc.* 82:1101–3
130. Marthaler, T. M. 1967. *Caries Res.* 1: 222–38
131. Newbrun, E. 1967. *Odontol. Revy* 18: 373–86
132. König, K. G. 1969. *Arch. Oral Biol.* 14: 991–93
133. Backer Dirks, O. 1974. *Caries Res.* 8: (Suppl. 1) 2–15
134. Ericsson, Y. 1974. *Caries Res.* 8:(Suppl. 1) 16–27
135. Aasenden, R., Peebles, T. C. 1974. *Arch. Oral Biol.* 19:321–26
136. Losee, F. L., Ludwig, T. G. 1970. *J. Dent. Res.* 49:1229–35
137. Büttner, W. 1969. *Caries Res.* 3:1–13
138. Manly, R. S. 1957. *J. Am. Coll. Dent.* 24:201–2
139. Jordan, H. V. 1973. *J. Can. Dent. Assoc.* 39:703–8
140. Handelman, S. L., Mills, J. R., Hawes, R. R. 1966. *J. Oral Ther.* 2:338–45
141. Englander, H. R., Keyes, P. H. 1971. *Arch. Oral Biol.* 16:469–72
142. DePaola, P. F., Jordan, H. V., Berg, J. 1974. *J. Dent. Res.* 53:108–14
143. Jordan, H. V., DePaola, P. F. 1974. *J. Dent. Res.* 53:115–20
144. Loesche, W. J., Murray, R. J., Mellberg, J. R. 1973. *Caries Res.* 7:283–96
145. Woods, R. 1971. *Aust. Dent. J.* 16: 152–55
146. Loesche, W. J., Nafe, D. 1973. *Arch. Oral Biol.* 18:1131–43
147. Löe, H., Schiott, C. R. 1970. *J. Periodont. Res.* 5:79–83
148. Gjermo, P., Baastad, K. L., Rölla, G. 1970. *J. Periodontal Res.* 5:102–9
149. Löe, H. 1970. In *Dental Plaque* ed. W. D. McHugh, 259–70. Edinburgh: Livingstone
150. Gaffar, A., Marcussen, H. W., Ruffner, J., Kestenbaum, R. C. 1970. *48th session of Int. Assoc. Dent. Res.*, p. 124 (Abstr.)
151. Bowen, W. H. 1969. *Brit. Dent. J.* 126:159–60

152. Tanzer, J. M., Hageage, G. L., Larson, R. H. 1973. *Arch. Oral Biol.* 18:1425–39
153. Taubman, M., Smith, D. J. 1974. *Infec. Immunity* 9:1079–91
154. Hayashi, J. A., Shklair, I. L., Bahn, A. N. 1972. *J. Dent. Res.* 51:436–42
155. Evans, R. T., Genco, R. J. 1973. *Infec. Immunity* 7:237–41
156. Olson, G. A., Bleiweiss, A. S., Small, P. A. 1972. *Infec. Immunity* 5:419–27
157. Olson, G. A., Guggenheim, B., Small, P. A. 1974. *Infec. Immunity* 9:273–78
158. Guggenheim, B. H., Mühlemann, H. R., Regolati, B., Schmid, R. 1970. In *Dental Plaque* ed. W. D. McHugh, 287–96. Edinburgh: Livingstone

Copyright 1975. All rights reserved

KETAMINE ANESTHESIA

❖7137

Charles F. Lanning, M.D. and Merel H. Harmel, M.D.
Department of Anesthesiology, Duke University Medical Center, Durham, North Carolina 27710

Ketamine [2-(*o*-chlorophenyl)-2-methylamino cyclohexanone HCl], a unique general anesthetic whose lack of cardiorespiratory depression is unequalled by any other general anesthetic currently available, was introduced in 1965 by McCarthy et al (1) and Domino, Chodoff & Corssen (2). It is an extremely versatile agent because it may be administered intravenously (i.v.) or intramuscularly (i.m.) without significant tissue irritation. As yet, contraindications to its use with other drugs used in anesthesia have not been demonstrated. Ketamine provides a profound analgesia characterized by a trance-like, cataleptic state with generalized mild hypertonus, nystagmus, cardiovascular stimulation, lack of respiratory depression, and laryngeal and pharyngeal competence. Onset of action is rapid, being 20–30 sec for i.v. and 3–5 min for i.m administration. Duration of surgical anesthesia is 5–10 min and 15–25 min following i.v. and i.m. doses. The usual range of ketamine dosage is 1–2 mg/kg i.v. and 5–10 mg/kg i.m. Maintenance doses are usually one half the induction dose. Its use for outpatient procedures has progressively decreased because time-to-ambulation may be prolonged, frequently lasting more than 2 hr. Wessels, Allen & Slogoff (3) have demonstrated that N_2O enhances ketamine by reducing the total dose of ketamine 39% and shortening the recovery period by 64%. Initially, ketamine did not appear to provide visceral analgesia; but, its use with N_2O and relaxants provides adequate anesthesia for intraabdominal and thoracic surgery (4). For these reasons, ketamine is now more commonly used in combination with other anesthetic agents rather than as a monoanesthetic.

Ketamine is rapidly distributed into all body tissues, primarily into fat, liver, lung, and brain. Metabolism occurs in the liver by N-demethylation and hydroxylation of the cyclohexanone ring with formation of water-soluble conjugates which are excreted in the urine. Human studies demonstrate an apparent biological half-life of 4 hr. with 70% urinary excretion in 24 hr. as the metabolized derivatives (5). However, Latarjet et al (6) have shown that ketamine's duration is not prolonged in humans with renal insufficiency.

Ketamine has proved to have a broad therapeutic index. McCarthy et al (1) demonstrated in dogs a ketamine LD_{50}/ED_{50} ratio five times that of pentobarbital. Repeated administrations have been given to humans without developing any significant tolerance or complication. Cumulation during anesthesia, however, does occur after multiple doses. Gross overdosage simply results in prolonged duration and respiratory depression. Postanesthetic anorexia (7) and immunosuppression (8), which are common to many general anesthetics, appear to be minimal with ketamine.

The pharyngeal and laryngeal reflexes remain active when ketamine is used as a monoanesthetic. Preservation of these reflexes, however, results in an increased incidence of laryngospasm, bronchospasm, and coughing secondary to secretions, blood, or manipulation in the oropharynx. These complications make ketamine a poor agent for endoscopy or oropharyngeal surgery. In addition, ketamine stimulates marked salivation which must be blocked by an antisialagogue. Despite retention of protective reflexes, tracheal soiling and aspiration may occur with ketamine; and, consequently, the airway must be safeguarded by one experienced in the management of airway obstruction and its complications. As with all general anesthetics, ketamine should be given with the usual precautions when a patient's stomach is full. Nevertheless, the degree of airway patency and lack of respiratory depression make ketamine the general anesthetic of choice when intubation is impossible. The decrease in airway resistance observed with ketamine makes it a desirable agent for bronchospastic and asthmatic patients (9).

Although the mechanism of ketamine's cardiovascular stimulation is unknown, detailed studies demonstrate an increase in mean aortic pressure, pulmonary artery pressure, central venous pressure, heart rate, and cardiac index, and a variable effect upon peripheral vascular resistance (10, 11). A summary of available information indicates that the adrenergic system must be intact in order for these responses to occur (12–14). Therefore, ketamine probably acts either directly by stimulating the central adrenergic centers or indirectly by inhibiting the neuronal uptake of catecholamines, especially norepinephrine (15). These cardiostimulatory responses, in addition to its antiarrhythmic properties (1, 16), make ketamine a good induction agent for poor risk and hypovolemic patients. However, for maintenance, ketamine may be a liability in patients with severe coronary insufficiency since myocardial oxygen consumption has been shown to be elevated by ketamine (10). In addition, dogs made hypotensive by bleeding develop a greater oxygen debt and deteriorate earlier with ketamine than with halothane or neuroleptanalgesia (17). After bronchopulmonary lavage in man, one of us (CFL) has noted a greater base-deficit with ketamine-relaxant anesthesia than with halothane-relaxant anesthesia despite comparable acid-base status during lavage.

The major deterrent to ketamine's wide acceptance has been the high incidence of postanesthetic emergence phenomena consisting of nausea, vomiting, restlessness, auditory and visual hallucinations, occasional frank delirium; and, especially, unpleasant dreams which have been reported to occur in 2–30% of patients (18, 19). Most of the dreams and hallucinatory phenomena may be related to misinterpretation of visual stimuli, resulting from nystagmus, and to a sense of bodily detachment which is probably created by a profound somatosensory analgesia. Although Hefez & Lanyi (20) have shown by reality testing that ketamine is not a psychotomimetic drug, the quality and content of the hallucinations and/or dreams may occasionally be compared to the reactions seen with common street hallucinogens.

Emergence phenomena occur more frequently with a high dose per unit operative time; and, consequently, they are more frequent with short operative procedures. They are uncommonly reported in children. For this reason, ketamine is more often used in children than adults; especially, in uncooperative children when an intravenous route is difficult to obtain. Emergence phenomena may be modified by narcotics and/or tranquilizers given as premedication or prior to the end of anesthesia (21–25). It has also been suggested that they may be reduced by allowing a patient to recover undisturbed (26). This, however, creates a logistic problem in busy recovery rooms.

The CNS effects of ketamine, as reflected in the EEG, have been interpreted by Corssen, Miyasaka, & Domino (26) as depression of the thalamo-neocortical system with activation of the limbic system. Because of this dual effect, they characterize ketamine as "dissociative anesthesia." Other authors (27–31), however, interpret the EEG changes solely as subclinical electroencephalographic seizure activity which they term "functional disorganization." For fear of creating frank convulsions, they recommend that ketamine be used with caution, if at all, in epileptic patients. In contrast, Corssen, Little & Tavakoli (32) found evidence to suggest that ketamine may have anticonvulsant properties and that it does not precipitate generalized convulsions in epileptic patients.

Ketamine generates an increase in CSF pressure, particularly in patients with abnormal CSF flow dynamics and/or other intracranial pathology (33). This increase may be attenuated by thiopental and controlled by induced hypocarbia. The effect upon the CSF appears to be directly related to an increase in cerebral blood flow produced by ketamine (34). In addition, cerebral oxygen consumption does not decrease as it does with other general anesthetics. Consequently, ketamine must be used with caution when a preexisting increase in CSF pressure is present.

Ketamine produces a predictable, but slight, rise in intraocular pressure. Some feel this elevation is not sufficient to contraindicate the use of ketamine for patients with glaucoma (7, 35). However, its use with an open globe is not recommended. In patients for ocular surgery, ketamine-induced nystagmus may be controlled with very low doses of muscle relaxants.

Providing no more than 2–3 mg/kg of ketamine is administered prior to umbilical cord clampage, significant newborn depression does not occur with vaginal or caesarean delivery (35, 36). With the limitations described, early fetal oxygenation and acid-base balance are unaltered (36, 37). An obstetrical advantage of ketamine is reduced uterine bleeding secondary to excellent postpartal uterine contractility (35, 38). Because of its long duration and the possibility of postanesthetic emergence phenomena, ketamine probably should not be used for spontaneous delivery; but, it may be beneficial for forceps delivery and obstetrical surgery such as in patients with placenta praevia. This objection to routine use may prove to be insignificant when ketamine is used in subanesthetic doses as recently described by Akamatsu et al (39).

In summary, ketamine is a general anesthetic which is in its third and last stage of clinical development. Now, as with other anesthetics, its beneficial qualities are being delineated and put to use for specific situations. Its major advantages are its ease of administration and rapid onset without cardiorespiratory depression, which make it particularly useful for poor risk patients and children. Ketamine's broad therapeutic index and lack of long term toxicity are indications for its use in situations such as repeated burn dressings and recurrent radiotherapy for uncooperative children. Its cardiovascular stimulation makes it a good induction agent for debilitated and hypovolemic patients. Its respiratory benefits make it useful for asthmatic patients and for patients in whom spontaneous ventilation is imperative. It is particularly useful for cardiac catheterization where myocardial depression and positive pressure ventilation may markedly change pressure and shunt abnormalities.

Ketamine is contraindicated in those circumstances where hypertension might be detrimental such as in patients with a history of cerebrovascular hemorrhage and intracranial, thoracic, and abdominal aneurysms. Marked hypertension may occur in patients receiving thyroxine medication (40). Corssen recommends that ketamine should not be used when systolic blood pressure is greater than 160 (26). This, however, is controversial because, in a hypertensive patient, hypotension may be even more detrimental than hypertension. Ketamine must be used with caution in patients with epilepsy or increased CSF pressure. It is not recommended for use in open globe eye surgery.

In conclusion, ketamine deserves a place in the anesthesiologist's armamentarium because of its many interesting features. However, future studies are needed to find means by which to control the major deterrents to its use; that is, postanesthetic emergence phenomena and ketamine-induced cardiovascular hyperreactivity.

Literature Cited

1. McCarthy, D. A., Chen, G., Kaump, D. H., Ensor, C. 1965. *J. New Drugs* 5:21–33
2. Domino, E. F., Chodoff, P., Corssen, G. 1965. *Clin. Pharmacol. Ther.* 6:279–91
3. Wessels, J. V., Allen, G. W., Slogoff, S. 1973. *Anesthesiology* 39:382–86
4. Vaughan, R. W., Stephen, C. R. 1974. *Anesth. Analg. (Cleveland)* 53:271–80
5. McCarthy, D. A., Glazko, A. J., Kurtz, S. June 18, 1970. *Medical Summary Ketalar CI-581,* 228–61. Ann Arbor, Michigan: Med. Sci. Aff. Div., Parke-Davis
6. Latarjet, J. et al 1972. *Anesth. Analg. Reanim.* 29:261–64
7. Dillon, J. B. 1971. *Proc. Roy. Soc. Med.* 64:1153–56
8. Cullen, B. F., Chretien, P. B. 1973. *Anesth. Analg. (Cleveland)* 52:518–21
9. Corssen, G., Gutierrez, J., Reves, J. G., Huber, F. C. Jr. 1972. *Anesth. Analg. (Cleveland)* 51:588–96
10. Tweed, W. A., Minuck, M., Mymin, D. 1972. *Anesthesiology* 37:613–19
11. Virtue, R. W. et al 1967. *Anesthesiology* 28:823–33
12. Chodoff, P. 1972. *Anesth. Analg. (Cleveland)* 51:247–50
13. Traber, D. L., Wilson, R. D., Priano, L. L. 1970. *Anesth. Analg. (Cleveland)* 49:420–26
14. Traber, D. L., Wilson, R. D. 1969. *Anesth. Analg. (Cleveland)* 48:248–52
15. Miletich, D. J., Ivankovic, A. D., Albrecht, R. F., Zahed, B., Ilahi, A. A. 1973. *Anesthesiology* 39:271–77
16. Dowdy, E. G., Kaya, K. 1968. *Anesthesiology* 29:931–43
17. Bruckner, J. B., Patschke, D., Reinecke, A., Tarnow, J. 1973. *Anaesthesiol. Wiederbelebung* 69:99–119
18. Albin, M. S., Dresner, A. J. June 18, 1970. *Medical Summary Ketalar CI-581,* 142–59. Ann Arbor, Michigan: Med. Sci. Aff. Div., Parke-Davis
19. Fine, J., Finestone, S. C. 1973. *Anesth. Analg. (Cleveland)* 52:428–30
20. Hefez, A., Lanyi, G. 1972. *Isr. Ann. Psychiat. Relat. Disciplines* 10:180–87

21. Boville, J. G., Clarke, R. S. J., Dundee, J. W., Pandit, S. K., Moore, J. 1971. *Brit. J. Anaesth.* 43:600–8
22. Becsey, L., Malamed, S., Radnay, P., Foldes, F. F. 1972. *Anesthesiology* 37: 536–42
23. Sadove, M. S. et al 1971. *Anesth. Analg.* (*Cleveland*) 50:526–32
24. Collier, B. 1972. *Anaesthesia* 27:120–34
25. Rita, L., Seleny, F. L. 1974. *Anesth. Analg.* (*Cleveland*) 53:380–82
26. Corssen, G., Miyasaka, M., Domino, E. F. 1968. *Anesth. Analg.* (*Cleveland*) 47:746–59
27. Mori, K., Kawamata, M., Mitani, H., Yamazaki, Y., Fujita, M. 1971. *Anesthesiology* 35:373–83
28. Kayama, K., Iwama, K. 1972. *Anesthesiology* 36:316–28
29. Winters, W. D., Ferrar-Allado, T., Guzman-Flores, C., Alcaraz, M. 1972. *Neuropharmacology* 11:303–15
30. Winters, W. D. 1972. *Anesthesiology* 36:309–12
31. Ferrer-Allado, T., Brechner, V. L., Dymond, A., Gozen, H., Crandall, P. 1973. *Anesthesiology* 38:333–44
32. Corssen, G., Little, S. C., Tavakoli, M. 1974. *Anesth. Analg.* (*Cleveland*) 319–35
33. Shapiro, H. M., Wyte, S. R., Harris, A. B. 1972. *Brit. J. Anaesth.* 44:1200–4
34. Takeshita, H., Okuda, Y., Sari, A. 1972. *Anesthesiology* 36:69–75
35. Langrehr, D., Neuhaus, R. 1973. *Acta Anaesthesiol. Belg.* 24:71–96
36. Galbert, M. W., Gardner, A. E. 1973. *Anesth. Analg.* (*Cleveland*) 52:926–30
37. Levinson, G., Shnider, S. M., Gildea, J. E., deLorimier, A. A. 1973. *Brit. J. Anaesth.* 45:1111–15
38. Galloon, S. 1973. *Can. Anaesth. Soc. J.* 20:141–45
39. Akamatsu, T. J., Bonica, J. J., Rehmet, R., Eng, M., Ueland, K. 1974. *Anesth. Analg.* (*Cleveland*) 53:284–87
40. Kaplan, J. A., Cooperman, L. H. 1971. *Anesthesiology* 35:229–30

Copyright 1975. All rights reserved

PATHOGENESIS OF HEPATIC COMA

◆7138

Leslie Zieve, M.D. and Demetre M. Nicoloff, M.D., Ph.D.
Departments of Medicine and Surgery, Minneapolis Veterans Hospital, University of Minnesota, Minneapolis, Minnesota 55417

Hepatic coma is a complex pathophysiologic state that is multifactorial in origin. Incomplete and uncontrolled but significant observations in patients have offered clues for experimental study. Unfortunately the body of solid experimental data is small, and the amount of uninhibited speculation excessive. To compound this problem there has been confusion between correlation and causation. Hepatic encephalopathy has many associated alterations, some of which are noncausal correlated changes and some consequences of the encephalopathy. Seen after the fact, these two types of alterations cannot be differentiated from causal factors. This has often been forgotten in this field where everyone is primarily interested in etiology, and claims have been made that do not withstand careful inspection.

Several previous reviews offer broad perspectives on the whole subject (1–3). In 1966 we (4) reviewed the subject of pathogenesis in detail, and we will assume acquaintance with that material for purposes of the present review. In addition to the known metabolic abnormalities, which will again be summarized and brought up to date, that article provided a brief discussion of the clinical picture, the role of precipitating factors, and the contribution of electrolyte and acid-base abnormalities. Since then Zimmon (5, 6) has shown that the well-known shift to the right in the in vivo oxyhemoglobin dissociation curve, when more than minimal, is associated with the presence of hepatic encephalopathy and is due to an unidentified small molecular weight acidic plasma factor. When the hepatic encephalopathy disappears, the shift to the right is reversed. Such reversal has been observed following suppression of gut bacteria.

KNOWN METABOLIC ABNORMALITIES

Brain Oxygen Consumption and Glucose Utilization Are Decreased

In acute hepatic coma of less than 24 hr duration brain oxygen consumption and glucose utilization are essentially normal (7). Therefore the onset of acute hepatic coma cannot be attributed to depression of brain energy metabolism. Beyond 24 hr of coma, oxygen consumption and glucose utilization decrease, the average reduction of each in samples of such patients being approximately 50% (7, 8). The correlation between the cerebral metabolic rate and the degree of neuropsychiatric involvement is poor (7). In some patients the mental state and the oxygen consumption may be influenced favorably by induction of metabolic alkalosis with sodium bicarbonate (9). The reason for this arousal effect is unknown.

Decreased utilization of oxygen characterizes all nonconvulsive (or preconvulsive) coma states (10). This is not due to lack of oxygen, but a reduced demand for oxygen. The adenylate energy charge, $([ATP] + \frac{1}{2}[ADP])/([ATP] + [ADP] + [AMP])$, was found to be normal in acute experimental hepatic coma and in other coma states due to fatty acids, mercaptans, barbiturates, or ether (11). It was also normal in rats made hyperammonemic with portacaval shunts (12). Thus these coma states do not disrupt the balance between production and utilization of energy in the brain. Oxidation is coupled normally to phosphorylation in acute experimental hepatic coma and other coma states (11). Therefore the reduced demand for oxygen must reflect decreased utilization of ATP.

Ammonia Is Increased in the Blood, and Ammonia and Glutamine Are Increased in Muscle, Brain, and Spinal Fluid

From clinical inconsistencies and from extensive experimental observations it seems clear that hepatic coma cannot be due simply to an excess of ammonia (13). We concluded previously that disturbed ammonia metabolism is in some way basic to the syndrome of hepatic coma, but that it is not the only etiologic abnormality (4). The primary disposition of ammonia in the brain is by the formation of glutamine. In rats with portacaval shunts a high correlation (0.92) was observed between blood ammonia and the brain content of glutamine (14). In man, Moore and associates (15) observed a high correlation ($r = 0.90$) between arterial and spinal fluid ammonia, but patients in hepatic coma could not be differentiated from those not in coma. In general, the correlation between blood ammonia and the degree of hepatic encephalopathy has been poor (4). This has been one of the main reasons there has been reluctance to accept ammonia toxicity as the primary etiologic agent in hepatic coma. In contrast there is a good correlation between spinal fluid glutamine concentration and the degree of encephalopathy (16). No other parameter has correlated as well with clinical severity.

α-Ketoglutaramate is derived from glutamine by removal of an amino group from the α-carbon. Like glutamine it was found increased in the spinal fluid of patients with hepatic coma (17). However the concentration of glutamine in these fluids was 42 times that of α-ketoglutaramate, suggesting that the latter as a minor metabolite probably has the same significance as glutamine. Coma was not produced when α-ketoglutaramate was perfused into the cerebral lateral ventricles of rats, and there is no reason to believe it has a role in pathogenesis of hepatic coma.

Several interesting experimental observations concerning ammonia toxicity have appeared whose implications have not been pursued. Nance & Kline (18) demonstrated in the germ-free Eck's fistula dog that portal-systemic encephalopathy occurred and was associated with hyperammonemia and increased numbers of brain astrocytes. The endogenous source of the increased ammonia was not established, but obviously gut bacteria are not essential to account for the toxic consequences of an Eck's fistula.

The well-known effects of glutamate and arginine in enhancing the disposition of ammonia have been discussed previously (4). Superstimulation of the urea cycle by a combination of carbamoyl glutamate and arginine protected rats completely from lethal doses of ammonium acetate (19). The combination was more effective than larger doses of arginine given alone. Similar protection against lethal doses of NH_4Cl was observed with methylglutamic acid in doses of 6 mmole per kg (20). This effect was not due to increased synthesis of urea nor of methylglutamine. In these experiments methionine sulfoximine was used to inhibit glutamine synthetase. Warren & Schenker (21) had previously shown that methionine sulfoximine protected mice against ammonia toxicity while the brain ammonia remained approximately twice normal. Unfortunately no measurements were made of brain glutamate and aspartate to see if they accumulated. The full implications of these observations have never been explored.

Pyruvate, Lactate, Citrate, and α-Ketoglutarate Are Increased in Blood and Muscle

The data in man have been cited previously (4). The blood levels reflect muscle and not the brain content of these metabolites. In the hepatectomized dog the increases in blood pyruvate, lactate, and α-ketoglutarate occurred within 30–60 min of removal of the liver, and, if anything, the concentrations decreased when the animals became comatose (22, 23). Thus the increased blood levels reflected the absence of the liver and not the presence of coma. In man, the elevated blood levels are a consequence of severe liver dysfunction, as is the occurrence of hepatic encephalopathy.

Ketone Production by the Liver and Blood Ketones Are Decreased

The work of Recant & Fischer was cited previously (4).

Fatty Acids Are Increased in Blood

Muto and Takahashi (24, 25) first focused on this effect of severe hepatic failure, reporting increases in the serum concentration of short-chain free fatty acids, particularly hexanoic acid (C_6). We (26) showed a 2-fold rise in the plasma volatile fatty acids ($<C_6$) as coma ensued following total hepatectomy in the dog. However the alterations in free fatty acids with hepatic failure are not limited to the short-chain fatty acids. An increase in the fasting plasma concentration of long-chain free fatty acids (Dole method) in severe hepatic failure was first noted by Stormont & Mackie (27) in 1959. Mortiaux & Dawson (28) found as much as a 2½-fold increase in the Dole plasma free fatty acid concentration in patients with liver disease; however, there was a broad range of variation in abnormalities observed. They could detect no apparent correlation between the total free fatty acid concentration and the degree of hepatic failure, the depth of jaundice, the serum albumin concentration,

or the neuropsychiatric complications. The flux of free fatty acids, i.e. amount disappearing from the plasma per minute, was increased about 40% in hepatic failure.

Mays (29) observed that the plasma free fatty acids per gram of serum albumin were increased about 12-fold in cirrhotics in coma, and only 4-fold in cirrhotics not in coma. Though based upon a small sample, his data are of interest because he found roughly similar increments in patients with posttraumatic delirium and pancreatic psychosis.

The plasma free fatty acids are of potential significance in relation to the possible toxicity of tryptophan, because the plasma free fatty acids have a role in determining the availability of tryptophan to the brain and thus have an effect on serotonin turnover (30, 31). They also have potential significance as toxic substances acting synergistically with other toxic substances, such as ammonia and mercaptans, which accumulate in hepatic failure (32, 33). These relationships will be discussed later.

Amino Acids Are Increased in Blood, Brain, Spinal Fluid and Urine

The early work on blood or urine aminoacids by Walshe, Bollman and co-workers, and Iber and colleagues has been cited previously (4). Bollman, Flock and associates (34, 35) had shown that in the hepatectomized or Eck's fistula dog the overwhelmingly predominant amino acid in brain and spinal fluid is glutamine. With the further development of chromatographic and ion-exchange column methods for separation of amino acids there has been a resurgence of interest in studying the abnormalities to be found in hepatic failure. In general, in the hepatectomized dog the straight-chain amino acids are increased in the plasma, and branched-chain amino acids (leucine, isoleucine, and valine) as well as their α-keto-analogues are decreased (22, 23). In the dog with massive acute hepatic necrosis a similar reduction in the plasma branched-chain aminoacids was observed and their clearance rates from plasma were increased (36). In contrast the clearance rates of methionine and phenylalanine were decreased and their concentrations in the plasma increased.

In patients with acute hepatic coma, consistently abnormal patterns of amino acids were observed in plasma and spinal fluid, with increased concentrations of methionine, glutamate or glutamine, aspartate or asparagine, phenylalanine, tyrosine, and histidine, and reduced concentrations of leucine, isoleucine, and valine (37).

Patients predisposed to development of hepatic encephalopathy when protein intake is normal have a consistent pattern of abnormality in plasma amino acids whether on a protein-restricted diet or receiving mixtures of amino acids by intravenous hyperalimentation (38). Plasma concentrations of methionine, phenylalanine, aspartate, and glutamate were increased by at least 50% and as much as 5-fold, and concentrations of leucine, isoleucine, and valine were decreased by approximately 40–50%.

In cirrhotics with a history of hepatic encephalopathy, amino acids have varying ammoniagenic potencies (39). Those with the greatest effect on the blood ammonia were threonine, serine, glycine, glutamine, histidine, lysine, and asparagine in that order. Those with the least ammoniagenic potency (less than 7% of that of glycine) were arginine, aspartate, glutamate, and tryptophan. The dietary source of protein influences the encephalopathic response of Eck's fistula patients (40) and dogs (41). Milk and cheese protein was better tolerated than meat protein. The meat fed to the dogs contained more than twice as much glycine, arginine, phenylalanine, and tryptophan as skim milk. There appears to be no consistency between these observations and the ammoniagenic potencies referred to above. The reason for the greater tolerance for milk proteins in patients predisposed to hepatic encephalopathy remains unknown.

A few recent studies have been made of amino acid alterations in the brain during experimental hepatic coma. In the hepatectomized dog and rat, tryptophan was increased 3- to 4-fold in the brain, 3- to 5-fold in heart or skeletal muscle, and 2-fold in plasma. Tyrosine was increased approximately 4- to 5-fold in brain, 3- to 6-fold in heart or skeletal muscle, and 6- to 9-fold in plasma (42). In contrast, brain 5-hydroxyindoleacetic acid (5HIAA) was increased only 2-fold and serotonin (5HT) was normal. The dog with acute massive hepatic necrosis showed marked elevations of brain tyrosine, phenylalanine, tryptophan, and histidine (43). Smaller but significant elevations of alanine, glycine, lysine, methionine, serine, and threonine were also observed. Aspartate and glutamate were decreased by 20–30%. Oddly enough these authors reported slight increases in the concentration of the branched-chain amino acids and made no mention at all of glutamine, so repetition of this work in this experimental model is needed.

In a similar experimental model of acute hepatic failure in pigs the abnormality in tryptophan was studied very carefully (44). The increase in brain tryptophan was verified and found to correlate highly ($r = 0.84$) with plasma free (but not total) tryptophan, which in turn correlated highly ($r = 0.90$) with plasma free fatty acids. From this and previous work by Curzon and associates (30, 31) it appears that in acute hepatic failure the increased turnover of brain serotonin, which is associated with increases in the concentration of brain tryptophan, is intimately dependent upon the plasma concentration of free fatty acids via the plasma free tryptophan. Besides the concentration of plasma free tryptophan (and plasma free fatty acids), another factor influencing the movement of tryptophan into the brain is its plasma concentration relative to five other amino acids (tyrosine, phenylalanine, leucine, isoleucine, and valine), which compete with it for uptake by the brain (45).

Tryptophan and tyrosine have special importance as precursors of neurotransmitters. However, tryptophan itself has toxic manifestations that have been confused with hepatic encephalopathy. Tryptophan is the most toxic amino acid when given parenterally to rats (46). Excess ingestion of this amino acid by normal man can cause drowsiness, euphoria, nystagmus, disturbances in gait, and altered reflexes (47). Doses of 3–4 g have been reported to cause drowsiness, dizziness, headache, and vomiting in most cirrhotics (48). In rhesus monkeys a high intake of tryptophan caused no neurologic abnormality (49). In an Eck's fistula dog, daily ingestion of as much as 10 g (1000 mg/kg) of tryptophan caused ataxia, visual disturbances, and incontinence (50). While this observation on a single dog, which was reported in a letter to the editor, has been cited repeatedly as an example of hepatic coma caused by tryptophan, it is clearly a different abnormality caused by a massive dose of tryptophan. It is of interest that "a large amount of meat or urea did not induce comparable symptoms despite a significant blood ammonia elevation in the same Eck's fistula dog" (50).

Three amino acids—aspartate, glutamate, and arginine—have particular theoretical, if not practical, significance for hepatic coma because they, or certain of their derivatives, have been shown to have arousal properties in patients with hepatic encephalopathy (4), and great potency in reversing experimental ammonia toxicity, as mentioned earlier. In the dog with acute massive hepatic necrosis brain aspartate and glutamate were reduced (43). It is of interest, though nothing was made of it at the time, that the chromatograms published in the early work by Flock and associates (34) on brain amino acids in the hepatectomized dog clearly show decreased amounts of aspartate in cerebrum, cerebellum, and muscle and of glutamate in cerebellum and muscle. After ammonia infusions, brain glutamine accumulated as in hepatic failure, however no change in the concentrations of aspartate and glutamate was observed (51). The whole problem of compartmentation of substances such as α-ketoglutarate, glutamate, and aspartate (as well as ATP) into pools of differing sizes and differing rates of turnover precludes the drawing of inferences from simple concentration measurements (51, 52).

Neurotransmitters Are Decreased in Brain and Muscle, Their Metabolites Are Increased in Spinal Fluid, and False Neurotransmitters Are Increased in Brain, Muscle, Blood, and Urine

We are belatedly recognizing that quantitatively glutamate and aspartate are probably the most important excitatory neurotransmitters (53). Only a small proportion of central nervous system synapses have transmission mediated by biogenic amines. Less than 5% of brain nerve terminals are mediated by catecholamines. Yet in relation to hepatic encephalopathy glutamate and aspartate as neurotransmitters have been entirely ignored while attention has focused on the biogenic amines.

In patients with hepatic encephalopathy due to fulminant hepatic failure or chronic liver disease, neurotransmitter metabolites such as glutamine, asparagine, homovanillic acid, and 5HIAA are increased in the spinal fluid, at times markedly so (37, 54). Their plasma precursors—glutamate, aspartate, phenylalanine, tyrosine, and free tryptophan—are increased as indicated previously. As in experimental acute hepatic failure in pigs, previously cited (44), the free tryptophan is strikingly high, presumably because of the increased free fatty acids observed (54). The false neurotransmitter, octopamine, is increased in blood and urine (55, 56). In such patients, L-dopa has an unpredictable but at times significant arousal effect (55, 57).

In the animal with a severely damaged or absent liver, the brain concentrations of aspartate, glutamate (equivocally), and norepinephrine are reduced (34, 43, 58, 59). Norepinephrine synthesis may be reduced (56). The concentration of serotonin has in most instances, and its derivative 5HIAA in all instances, been found elevated (42, 44, 58). The alterations in norepinephrine and serotonin have been reversed by L-dopa (58). In rats with ammonia-induced coma these changes in brain norepinephrine and serotonin were not observed (60).

False neurotransmitters are also increased in brain, muscle, and blood of animals with acute massive hepatic necrosis leading to coma (55, 59). This has similarly been observed in Eck's fistula rats, in which intestinal sterilization decreased but did not entirely reverse the elevated brain octopamine levels (58). L-Dopa has a similar effect on the elevated brain octopamine as well as blood ammonia (3).

After administration of a monoamine oxidase inhibitor, the concentration of octopamine in tissues rises severalfold. The result is central stimulation rather than coma. Feeding of phenylethylamine or of phenylalanine to Eck's fistula rats results in a 2½-fold increase in brain octopamine within 2 hr (61). This does not apparently produce coma since no such effect was mentioned. We have instilled octopamine into the cerebral ventricle of rats without causing coma. One can hardly attribute hepatic coma to a direct toxic effect of the false neurotransmitters themselves.

Mercaptans Are Increased in Breath, Blood, and Urine

Challenger & Walshe (62) established the association of mercaptans and hepatic failure by isolating methanethiol (methyl mercaptan) from the urine of a woman in severe acute hepatic coma. They were interested in the cause of fetor hepaticus and concluded that it was probably due to pulmonary excretion of a mixture of methanethiol, dimethyl sulfide, and dimethyl disulfide. This probably valid inference has never been established by breath analyses, however we (63) have shown that (*a*) mercaptans are normally present in the breath, (*b*) in hepatic failure breath methanethiol increases approximately 4-fold, (*c*) the highest values are observed in cirrhotics with gastrointestinal bleeding, and (*d*) cirrhotics fed methionine have a distinctive fetor, related to but not identical with fetor hepaticus, which is due to dimethyl sulfide, the primary derivative of methanethiol.

Mercaptans cause reversible coma in animals, and enhance the toxicity of ammonia and fatty acids (33). We have suggested that it may play a role as an augmenter or intensifier of the toxic effects of ammonia and fatty acids in human hepatic failure.

In terms of biochemical pharmacology, practically nothing is known about the toxicology of this class of compounds. Ahmed and co-workers (64) have initiated such studies by finding that methanethiol at low concentrations inhibits rat brain Na^+, K^+-ATPase, and that this effect is freely reversible. The effect of methanethiol differs from the known inhibitory effect of free fatty acids on this important membrane enzyme (65).

POSSIBLE ROLE OF REDUCED CELLULAR OXIDATIVE METABOLISM

In Figure 1 we have brought together in the context of cell metabolism all the known abnormalities that have been observed in blood or spinal fluid of patients with hepatic encephalopathy. We had concluded previously (4) that reduced cellular oxidative metabolism was probably the fundamental basis for the coma of liver failure. At the time this seemed a reasonable hypothesis, given the data available,

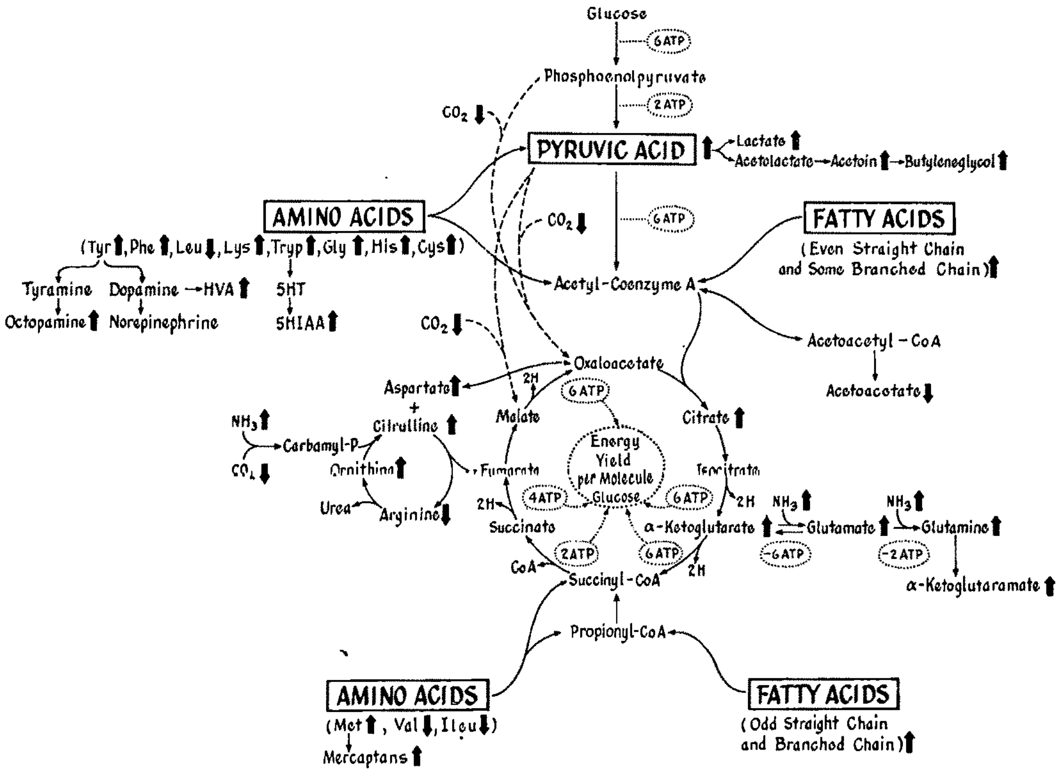


Figure 1 Metabolic map of abnormalities observed in blood and spinal fluid of patients with hepatic encephalopathy. Heavy arrows indicate items that have been measured and the direction of change.

since the most striking metabolic event that correlated with the occurrence of hepatic encephalopathy was the accumulation of glutamine in the brain and spinal fluid. This prompted us to look for a unique abnormality in cerebral glucose metabolism in the hepatectomized dog in coma, which we did not find. As shown in Figure 2, there was the expected increase in brain ammonia and glutamine, but there was no significant change from controls in the brain content of the glycolytic substrates measured, nor in the ATP concentration. There was a significant reduction in the citric acid cycle substrates α-ketoglutarate, fumarate, malate, and oxaloacetate, but this was no different than that observed by simple sedation with Innovar®, an analgesia-tranquilizer combination that did not result in the loss of consciousness. In the mouse, Goldberg and associates (66) showed essentially the same effect on the citric acid cycle with ether, amytal, and phenobarbital. Thus there was nothing unique about experimental hepatic coma with respect to cerebral glucose metabolism, the alterations observed being similar to those produced by administration of external agents that result in sedation or anesthesia.

The Bessmans (67) had suggested that hepatic coma might be due to depletion of α-ketoglutarate in the brain due to its diversion from the Kreb's cycle into the pathway toward glutamine as a result of the excess ammonia. Clarke (68) has pointed out that on the basis of our knowledge of the rate of turnover of citric acid cycle substrates, the duration of survival of the animal or patient in hepatic coma is inconsistent with the hypothesis that depletion of endogenous α-ketoglutarate is the cause of the coma. Thus on the face of it, the Bessman hypothesis is untenable. The same argument applies to the hypothesis that the coma is due to depletion of ATP. The existence of pools of each of these substances, of varying size and varying turnover rates, compounds the problem.

Much work (69–71) has been directed at alterations in concentrations of high energy phosphates (ATP and phosphocreatine) in pure ammonia-induced coma in rats, which has little direct relevance to the problem of hepatic coma for several reasons. First, acute ammonia-induced coma is a state of extreme hyperkinesis in which the animals are subconvulsive (or convulsive) and the turnover of metabolites is rapid. Oxidative metabolism is increased (72). This is not ordinarily the case in either human or experimental hepatic coma. Second, tissue hypoxia of some degree is unavoidable during the process of freezing the tissue and identifying and separating the component parts, such as the brain stem. This results in variable loss of the extremely labile metabolites and artifactual decreases in concentration (73). This may have contributed to inconsistencies that have been observed (69, 70). Third, concentration measurements, even if they were valid, give no information regarding turnover of high energy phosphates, which is the relevant information for understanding the alterations in energy metabolism. As in other coma states, the balance between production and utilization of energy by the brain is probably maintained up to the point of death, and one would anticipate that the energy charge of the adenylate pool in ammonia-induced coma would be normal if tissue hypoxia were avoided entirely. A contrary viewpoint regarding the significance of these studies will be found in reference 2.

From the present vantage point it seems less likely that reduced cellular oxidative metabolism will provide an explanation for the coma of hepatic failure, even though

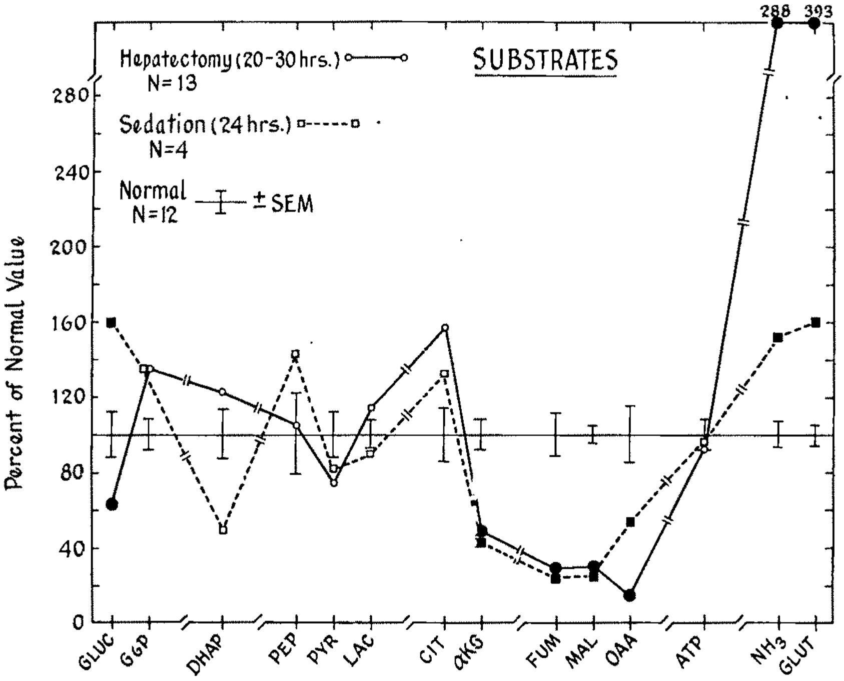


Figure 2 Relative cerebral alterations in selected glycolytic pathway and citric acid cycle substrates of dogs in coma following one-stage total hepatectomy. The substances measured in brain were: Gluc, glucose; G6P, glucose-6-phosphate; DHAP, dihydroxyacetone phosphate; PEP, phosphoenolpyruvate; PYR, pyruvate; LAC, lactate; CIT, citrate; αKG, α-ketoglutarate; FUM, fumarate; MAL, malate; OAA, oxaloacetate; ATP, adenosinetriphosphate; NH_3, ammonia; GLUT, glutamine. The solid line and circles represent dogs in coma 20–30 hr after total hepatectomy. These animals received i.v. glucose continuously to sustain blood glucose above 60 mg/100 ml. The broken line and squares indicate dogs sedated for 24 hr with Innovar® [fentanyl (analgesic) + droperidol (tranquilizer)]. Closed circles and squares indicate statistically significant deviations from normal.

abnormalities in oxidative metabolism are demonstrable. It would appear desirable to take a fresh look at the role of toxic factors, in particular their effects on membrane function and nerve cell transmission as well as their cumulative effects on cell metabolism.

POSSIBLE IMPORTANCE OF INTERACTIONS AMONG TOXIC FACTORS

The only toxic substances that are known to accumulate abnormally in hepatic failure and also cause reversible coma in experimental animals are ammonia, fatty acids, and mercaptans (4, 33). Substances like α-ketoglutaramate, certain biogenic amines, and tryptophan, for which etiologic claims have been made, do not have this critical property. In experimental animals the coma-producing effects of ammonia, fatty acids, and mercaptans have been shown to be strikingly interdependent (32, 33). Blood levels of these substances at which coma occurs in normal animals are markedly reduced when any two of them are injected simultaneously. The dose of ammonium salt required to produce coma in normal animals is sharply reduced by the presence of a subcoma dose of either a fatty acid or mercaptan. The same effect on the dose-response curve to octanoic acid is observed in the presence of subcoma doses of either NH_4^+ or mercaptan. The incidence of coma can be raised from 0 to 100% by injection of subcoma doses of any two of these substances simultaneously.

The blood ammonia concentration resulting from the injection of NH_4^+ is also markedly affected by the simultaneous injection of subcoma doses of either a fatty acid or a mercaptan. In rats, for example, the average blood ammonia rose from 1042 to 2585 μg/100 ml when the maximum non-coma-producing doses of NH_4^+ and sodium octanoate were injected simultaneously, and the animals then became comatose. This effect of fatty acids on blood ammonia can be demonstrated with long-chain as well as short-chain fatty acids (32), and with dimethyl sulfide as well as methanethiol (33). An exogenous coma-producing substance such as pentobarbital, not primarily related to liver failure, does not have this synergistic effect on blood ammonia, though it does have an additive effect on the production of coma. The effect of fatty acids on blood ammonia can be related to their inhibitory effects on urea synthesis and glutamate dehydrogenase activity (Zieve, unpublished observations). A similar converse effect of NH_4^+ and fatty acids on the blood methanethiol has also been observed (33).

In rats and cats with acute experimental hepatic coma resulting from massive liver necrosis, there is a high curvilinear correlation ratio (0.92) between blood ammonia and plasma free fatty acids (32). The average blood levels of ammonia, free fatty acids, and methanethiol in such rats with hepatic coma are much lower than blood levels required to produce coma in normal rats with each of these substances individually, that is, ammonia 1480 vs 2300 μg/100 ml, free fatty acids 490 vs 2200 μeq/liter, methanethiol ≤ 0.1 vs 0.5 nmole/ml. When normal rats are given a combination of doses of NH_4^+, fatty acid, and mercaptan which produce blood levels in the range of those observed in rats with experimental hepatic coma, they become comatose. Thus the synergistic action of these three toxic substances seems sufficient to account for the coma of acute experimental hepatic necrosis (33). The mercaptan appears to play a role as an augmenter or intensifier of the toxic effects of ammonia particularly and to some extent of fatty acids.

Measurements of a similar sort in blood or spinal fluid of patients with hepatic failure are not available at present. The evidence of synergism in animals suggests that similar interrelationships should be sought among variables reflecting the effects of suspected toxins in patients with hepatic encephalopathy. The multifactorial nature of hepatic coma, which has been accepted conceptually, requires a multifactorial approach to measurements in order to recognize and allow for interdependence and synergistic effects.

CONCLUSIONS AND IMPLICATIONS

The relative importance of the presence of toxic factors as against the lack of essential substances in the pathogenesis of hepatic encephalopathy in man has never been explored. It seems quite likely that both are important. The foregoing observations in experimental animals have, perhaps, a direct implication for human hepatic coma that follows fulminant viral or toxic hepatitis with acute massive necrosis. However the usual variety of human hepatic encephalopathy that is an end result of chronic hepatic failure and progressive over a long period of time, is undoubtedly more complex and involves pathophysiologic alterations that intensify or augment the narcotic effects of primary toxins when they are present in small amounts. For example, hypoalbuminemia would be expected to exaggerate the toxic effects of fatty acids because of the reduced availability of protein binding sites for fatty acids and the greater potentiality for movement of free fatty acids into cells.

In a patient with severe liver disease, the presence and severity of encephalopathy will, we think, depend upon a balance between coma-producing substances and coma-preventing substances. The relative significance of any single substance will presumably vary from case to case. Preeminent among the coma-producers would be ammonia, fatty acids, mercaptans, and possibly certain amines. Among the coma-preventers one might consider, depending upon circumstances, electrolytes, albumin, L-dopa, and certain amino acids (arginine, aspartate, glutamate). Though there are many unsettled questions in this area, it seems safe to conclude that alterations in a number of toxic as well as protective factors are involved in the pathogenesis of hepatic coma, and that a variable quantitative balance between the combined effects of coma-producers and coma-preventers is involved. The precise locus of action of the coma-producers is unknown. The fragmentary experimental data that are available suggest that any one or more of three loci—the nerve cell membrane, cellular oxidative metabolism, and neurotransmission—may be the critical focal point. Similarities have been observed between the effects of anesthetics (74, 75) and the effects of experimental hepatic coma, or of a fatty acid or mercaptan coma on oxygen utilization, oxidative metabolism, and membrane action. We suspect therefore that, in terms of brain mechanisms, the coma of hepatic failure will be found to have much in common with ordinary anesthesia (64, 76).

Literature Cited

1. Gabuzda, G. J. 1962. Hepatic coma: Clinical considerations, pathogenesis, and management. *Advan. Intern. Med.* 11:11–73
2. Schenker, S., Breen, K. J., Hoyumpa, A. M. 1974. Hepatic encephalopathy: Current status. *Gastroenterology* 66: 121–51
3. Fischer, J. E. 1974. Hepatic coma in cirrhosis, portal hypertension, and following portacaval shunt. *Arch. Surg.* 108:325–36
4. Zieve, L. 1966. Pathogenesis of hepatic coma. *Arch. Intern. Med.* 118:211–23
5. Zimmon, D. S. 1967. Oxyhemoglobin dissociation in patients with hepatic encephalopathy. *Gastroenterology* 52: 547–54
6. Zimmon, D. S. 1972. Changing oxyhemoglobin dissociation curve in hepatic encephalopathy. *Clin. Res.* 20:469 (Abstr.)
7. Maiolo, A. T., Porro, G. B., Galli, C., Sessa, M., Polli, E. E. 1971. Brain en-

ergy metabolism in hepatic coma. *Exp. Biol. Med.* 4:52–70

8. Fazekas, J. F., Ticktin, H. E., Ehrmantraut, W. R., Alman, R. W. 1956. Cerebral metabolism in hepatic insufficiency. *Am. J. Med.* 21:843–49
9. James, I. M., Sampson, D., Nashat, S., Williams, H. S., Garassini, M. 1969. Effect of induced metabolic alkalosis in hepatic encephalopathy. *Lancet* 2: 1106–8
10. Lassen, N. A. 1959. Cerebral blood flow and oxygen consumption in man. *Physiol. Rev.* 39:183–238
11. Derr, R. F., Zieve, L. 1973. Decreased cerebral uptake of oxygen in coma—a consequence of decreased utilization of ATP. *J. Neurochem.* 21:1555–57
12. Holmin, T., Siesjö, B. K. 1974. The effect of porta-caval anastomosis upon the energy state and upon acid-base parameters of the rat brain. *J. Neurochem.* 22:403–12.
13. Walker, C. O., Schenker, S. 1970. Pathogenesis of hepatic encephalopathy with special reference to the role of ammonia. *Am. J. Clin. Nutr.* 23:619–32
14. Williams, A. H., Kyu, M. H., Fenton, J. C. B., Cavanagh, J. B. 1972. The glutamate and glutamine content of rat brain after portocaval anastomosis. *J. Neurochem.* 19:1073–77
15. Moore, E. W., Strohmeyer, G. W., Chalmers, T. C. 1963 Distribution of ammonia across the blood-cerebrospinal fluid barrier in patients with hepatic failure. *Am. J. Med.* 35:350–62
16. Hourani, B. T., Hamlin, E. M., Reynolds, T. B. 1971. Cerebrospinal fluid glutamine as a measure of hepatic encephalopathy. *Arch. Intern. Med.* 127: 1033–36
17. Vergara, F., Plum, F., Duffy, T. E. 1974. α-Ketoglutaramate: Increased concentrations in the cerebrospinal fluid of patients in hepatic coma. *Science* 183:81–83
18. Nance, F. C., Kline, D. G. 1971. Eck's fistula encephalopathy in germfree dogs. *Ann. Surg.* 174:856–62
19. Kim, S., Paik, W. K., Cohen, P. P. 1972. Ammonia intoxication in rats: Protection by *N*-carbamoyl-L-glutamate plus L-arginine. *Proc. Nat. Acad. Sci. USA* 69:3530–33
20. Lamar, C. Jr. 1970. Ammonia toxicity in rats: Protection by α-methylglutamic acid. *Toxicol. Appl. Pharmacol.* 17:795–803
21. Warren, K. S., Schenker, S. 1964. Effect of an inhibitor of glutamine synthesis (Methionine sulfoximine) on ammonia toxicity and metabolism. *J. Lab. Clin. Med.* 64:442–49
22. McMenamy, R. H., Vang, J., Drapanas, T. 1965. Amino acid and α-keto acid concentrations in plasma and blood of the liverless dog. *Am. J. Physiol.* 209: 1046–52
23. Drapanas, T., McMenamy, R. H., Adler, W. J., Vang, J. O. 1965. Intermediary metabolism following hepatectomy in dogs. *Ann. Surg.* 162:621–33
24. Muto, Y., Takahashi, Y. 1964. Gas chromatographic determination of plasma short chain fatty acids in diseases of the liver. *J. Jap. Soc. Intern. Med.* 53:828–39
25. Muto, Y. 1966. Clinical study on the relationship of short chain fatty acids and hepatic encephalopathy. *Jap. J. Gastroenterol. 63:19–32*
26. Zieve, L., Nicoloff, D. M., Mahadevan, V. 1968. Effect of hepatectomy on volatile free fatty acids of the blood. *Gastroenterology* 54:1285 (Abstr.)
27. Stormont, J. M., Mackie, J. E. 1959. An abnormality of nonesterified fatty acid metabolism in alcoholic patients with cirrhosis. *J. Clin. Invest.* 38:1047 (Abstr.)
28. Mortiaux, A., Dawson, A. M. 1961. Plasma free fatty acid in liver disease. *Gut* 2:304–9
29. Mays, E. T. 1972. Encephalopathy and fatty acid toxicity. *Surg. Forum* 23: 352–54
30. Knott, P. J., Curzon, G. 1972. Free tryptophan in plasma and brain tryptophan metabolism. *Nature* 239:452–53
31. Curzon, G., Friedel, J., Knott, P. J. 1973. The effect of fatty acids on the binding of tryptophan to plasma protein. *Nature* 242:198–200
32. Zieve, F. J., Zieve, L., Doizaki, W. M., Gilsdorf, R. B. 1974. Synergism between ammonia and fatty acids in the production of coma: Implications for hepatic coma. *J. Pharmacol. Exp. Ther.* 191:10–16
33. Zieve, L., Doizaki, W. M., Zieve, F. J. 1974. Synergism between mercaptans and ammonia or fatty acids in the production of coma: A possible role for mercaptans in the pathogenesis of hepatic coma. *J. Lab. Clin. Med.* 83:16–28
34. Flock, E. V., Block, M. A., Grindlay, J. H., Mann, F. C., Bollman, J. L. 1953. Changes in free amino acids of brain and muscle after total hepatectomy. *J. Biol. Chem.* 200:529–36

35. Bollman, J. L., Flock, E. V., Grindlay, J. H., Bickford, R. G., Lichtenheld, F. R. 1957. Coma with increased amino acids of brain and cerebrospinal fluid in dogs with Eck's fistula. *Arch. Surg.* 75:405–12
36. Iob, V. et al 1970. Alterations in plasma-free amino acids in dogs with hepatic insufficiency. *Surg. Gynecol. Obstet.* 130:794–800
37. Knell, A. J., Curzon, G., Pratt, O. E., Williams, R. S. 1972. Neurotransmitter and aminoacid metabolism in acute hepatic coma. *Proc. Int. Assoc. Study of the Liver* 5:40 (Abstr.)
38. Fischer, J. E. et al 1974. Plasma amino acids in patients with hepatic encephalopathy. *Am. J. Surg.* 127:40–47
39. Rudman, D., Galambos, J. T., Smith, R. B. III, Salam, A. A., Warren, W. D. 1973. Comparison of the effect of various amino acids upon the blood ammonia concentration of patients with liver disease. *Am. J. Clin. Nutr.* 26: 916–25
40. Fenton, J. C. B., Knight, E. J., Humpherson, P. L. 1966. Milk-and-cheese diet in portal-systemic encephalopathy. *Lancet* 1:164–66
41. Condon, R. E. 1971. Effect of dietary protein on symptoms and survival in dogs with an Eck fistula. *Am. J. Surg.* 121:107–13
42. Tyce, G. M., Flock, E. V., Owen, C. A. Jr., Stobie, G. H. C., David, C. 1967. 5-Hydroxyindole metabolism in the brain after hepatectomy. *Biochem. Pharmacol.* 16:979–92
43. Mattson, W. J. Jr. et al 1968. Brain amino acids in acute hepatic coma. *Surg. Forum* 19:331–32
44. Curzon, G. et al 1973. Plasma and brain tryptophan changes in experimental acute hepatic failure. *J. Neurochem.* 21:137–45
45. Fernstrom, J. D., Wurtman, R. J. 1972. Brain serotonin content: Physiological regulation by plasma neutral amino acids. *Science* 178:414–16
46. Guillino, P. et al 1956. Studies on the metabolism of amino acids and related compounds in vivo. I. Toxicity of essential amino acids, individually and in mixtures, and the protective effect of L-arginine. *Arch. Biochem. Biophys.* 64:319–32
47. Smith, B., Prockop, D. J. 1962. Central-nervous-system effects of ingestion of L-tryptophan by normal subjects. *N. Engl. J. Med.* 267:1338–41
48. Hirayama, C. 1971. Tryptophan metabolism in liver disease. *Clin. Chim. Acta* 32:191–97
49. Curzon, G., Ettlinger, G., Cole, M., Walsh, J. 1963. The biochemical, behavioral, and neurologic effects of high L-tryptophan intake in the Rhesus monkey. *Neurology* 13:431–38
50. Ogihara, K., Mozai, T., Hirai, S. 1966. Tryptophan as cause of hepatic coma. *N. Engl. J. Med.* 275:1255–56
51. Berl, S., Takagaki, G., Clarke, D. D., Waelsch, H. 1962. Metabolic compartments in vivo. Ammonia and glutamic acid metabolism in brain and liver. *J. Biol. Chem.* 237:2562–69
52. Berl, S., Clarke, D. D. 1969. Compartmentation of amino acid metabolism. *Handbook of Neurochemistry, Vol. II. Structural Neurochemistry,* ed. A. Lajtha, Chap. 19, 447–72. New York: Plenum
53. Snyder, S. H., Young, A. B., Bennett, J. P., Mulder, A. H. 1973. Synaptic biochemistry of amino acids. *Fed. Proc.* 32:2039–47
54. Knell, A. J., Davidson, A. R., Williams, R., Kantamaneni, B. D., Curzon, G. 1974. Dopamine and serotonin metabolism in hepatic encephalopathy. *Brit. Med. J.* 1:549–51
55. Fischer, J. E., Baldessarini, R. J. 1971. False neurotransmitters and hepatic failure. *Lancet* 1:75–80
56. Lam, K. C., Tall, A. R., Goldstein, G. B., Mistilis, S. P. 1973. Role of a false neurotransmitter, octopamine, in the pathogenesis of hepatic and renal encephalopathy. *Scand. J. Gastroenterol.* 8:465–72
57. Parkes, J. D., Sharpstone, P., Williams, R. 1970. Levodopa in hepatic coma. *Lancet* 2:1341–43
58. Fischer, J. E., James, J. H., Baldessarini, R. 1972. Changes in brain amines following portal flow diversion and acute hepatic coma: Effects of levodopa (L-dopa) and intestinal sterilization. *Surg. Forum* 23:348–50
59. Dodsworth, J. M., James, J. H., Cummings, M. C., Fischer, J. E. 1974. Depletion of brain norepinephrine in acute hepatic coma. *Surgery* 75:811–20
60. Walker, C. O., Speeg, K. V. Jr., Levinson, J. D., Schenker, S. 1971. Cerebral acetylcholine, serotonin, and norepinephrine in acute ammonia intoxication. *Proc. Soc. Exp. Biol. Med.* 136: 568–71
61. Fischer, J. E., James, J. H. 1971. Mech-

anism of action of L-dopa in hepatic coma. *Surg. Forum* 22:347–49

62. Challenger, F., Walshe, J. M. 1955. Foetor hepaticus. *Lancet* 1:1239–41
63. Chen, S., Zieve, L., Mahadevan, V. 1970. Mercaptans and dimethyl sulfide in the breath of patients with cirrhosis of the liver. *J. Clin. Lab. Med.* 75: 628–35
64. Foster, D., Ahmed, K., Zieve, L. 1974. Action of methanethiol on Na^+, K^+-ATPase: Implications for hepatic coma. *Ann. N. Y. Acad. Sci.* In press
65. Ahmed, K., Thomas, B. S. 1971. The effects of long chain fatty acids on sodium plus potassium ion-stimulated adenosine triphosphatase of rat brain. *J. Biol. Chem.* 246:103–9
66. Goldberg, N. D., Passonneau, J. V., Lowry, O. H. 1966. Effects of changes in brain metabolism on the levels of citric acid cycle intermediates. *J. Biol. Chem.* 241:3997–4003
67. Bessman, S. P., Bessman, A. N. 1955. The cerebral and peripheral uptake of ammonia in liver disease with an hypothesis for the mechanism of hepatic coma. *J. Clin. Invest.* 34:622–28
68. Clarke, D. D. 1971. Comments on cerebral aminoacid metabolism in hepatic coma. *Exp. Biol. Med.* 4:79–84
69. Schenker, S., McCandless, D. W., Brophy, E., Lewis, M. S. 1967. Studies on the intracerebral toxicity of ammonia. *J. Clin. Invest.* 46:838–48
70. Hindfelt, B., Siesjö, B. K. 1971. Cerebral effects of acute ammonia intoxication. II. The effect upon energy metabolism. *Scand. J. Clin. Lab. Invest.* 28: 365–74
71. Hindfelt, B. 1972. The effect of sustained hyperammonemia upon the metabolic state of the brain. *Scand. J. Clin. Lab. Invest.* 30:245–55
72. Hawkins, R. A., Miller, A. L., Nielsen, R. C., Veech, R. L. 1973. The acute action of ammonia on rat brain metabolism in vivo. *Biochem. J.* 134:1001–8
73. Zieve, F. J., Zieve, L., Gilsdorf, R. B. 1972. Brain ATP concentrations in hepatic coma. *Gastroenterology* 62:880 (Abstr.)
74. Cohen, P. J. 1972. The metabolic function of oxygen and biochemical lesions of hypoxia. *Anesthesiology* 37:148–77
75. Seeman, P. 1972. The membrane actions of anesthetics and tranquilizers. *Pharmacol. Rev.* 24:583–655
76. Quarfoth, G., Foster, D., Ahmed, K., Zieve, L. 1974. Action of methanethiol on erythrocyte membranes and on brain Na^+, K^+-ATPase: Implications in hepatic coma. *Pharmacologist* 16:293 (Abstr.)

Copyright 1975. All rights reserved

CLINICAL PHARMACOLOGY OF DIGITALIS GLYCOSIDES[1]

❖7139

James E. Doherty, M.D., F.A.C.P., F.A.C.C. and James J. Kane, M.D.
Division of Cardiology, Department of Medicine, Veterans Administration Hospital, and
University of Arkansas School of Medicine, Little Rock, Arkansas 72206

INTRODUCTION

The clinical importance of digitalis glycosides in American medicine is illustrated by the fact that in 1971 digoxin was fourth on the list of most frequently prescribed drugs; digitoxin, sixteenth; and digitalis leaf, nineteenth (1). The first botanical with documented inotropic properties, the squill (urginea maritima or sea onion), was used by Egyptian physicians about 1500 B.C. (2). Continuing scientific interest is demonstrated by the listing of more than 600 citations for digitalis and similar products in the 1972 *Index Medicus* (3). The purpose of this review will be to apply appropriate clinical pharmacologic principles to these very commonly used preparations which possess the narrowest therapeutic:toxic ratio of almost any drug in common use today.

PHARMACODYNAMICS OF DIGITALIS PREPARATIONS

The pharmacodynamic properties of digitalis glycosides include the following:

1. Inotropic effect—the induction of increased increments of contractile force of cardiac muscle.
2. Chronotropic effect—the "slowing" effect on heart rate.
3. Toxic and/or side effects—undesirable reactions to the drugs brought about by excessive concentration of the glycoside at the myocardial receptor site and/or extracardiac binding sites. Idiosyncracy to digitalis, although rare, may also be included here.

Inotropic Effect of Digitalis Glycosides

The ability of the glycoside to cause increased myocardial fiber shortening is presumably related to inhibition of the sodium-potassium ATPase enzyme which sec-

---

[1] Supported in part by National Institutes of Health Grant 5R01 HL 06642–12 and Veterans Administration Medical Research Information System Grant 1887–01.

ondarily alters the amount of calcium ion available for binding to specific myocardial proteins (4). The recent work of Okita (5) questions the role of Na-K ATPase in this area and it may be hoped that further work will resolve this controversy.

Chronotropic Effect of Digitalis Glycosides

The chronotropic or "slowing" effect of digitalis glycosides is thought to have two phases: 1. a "vagal" effect which is operative very shortly after administration of the drug, and 2. a direct myocardial effect which relates to (*a*) improved cardiac performance in the intact subject and (*b*) slowing of impulse propagation from the atria to the ventricles in the atrio-ventricular nodal area (6).

Toxic and/or Side Effects of Digitalis Glycosides

Digitalis intoxication as a clinical entity seems to be increasing in incidence in the hospitalized cardiac patient. It probably represents the most frequent adverse drug reaction (as well as the most dangerous) physicians observe. A full scale discussion of digitalis toxicity will follow the section on pharmacokinetics of the various glycosides, as duration of intoxication is dependent on kinetic properties of the various preparations.

PHARMACOKINETICS OF DIGITALIS GLYCOSIDES

Much of the recent scientific data regarding the kinetic behavior of digitalis glycosides has been generated through turnover studies with a radioactive label (^{14}C, ^{3}H) applied to the compound. These studies are supplemented by later studies of serum levels and excretory products by inhibitions of red cell rubidium uptake by the sample (7) or radioimmunoassay (8).

Kinetic behavior is well described for the purified glycosides: digoxin, digitoxin, ouabain, deslanoside, and the similar product lanatoside C. Digitalis leaf is a mixture of glycosides containing mostly digitoxin and is probably closest in its kinetic behavior to that glycoside. An attempt will be made to discuss in depth the kinetics of digoxin and digitoxin.

Digoxin Pharmacokinetics

Digoxin (12-hydroxydigitoxin) is now the most popular glycoside used in the United States by both oral and parenteral routes of administration.

DIGOXIN TURNOVER KINETICS AFTER ORAL ADMINISTRATION Tritium-labeled digoxin was given by mouth to 12 human subjects who were in congestive heart failure. Digoxin was recovered from serum, urine, and stool for a 7-day period and digoxin extracted by a technique previously reported (9). Figure 1 is a composite graph of the serum turnover of these patients. Radioactivity representing digoxin appears in the serum almost immediately and rises to a peak level at 45–60 min and then falls more gradually to a plateau level at 5–6 hr after administration of the drug.

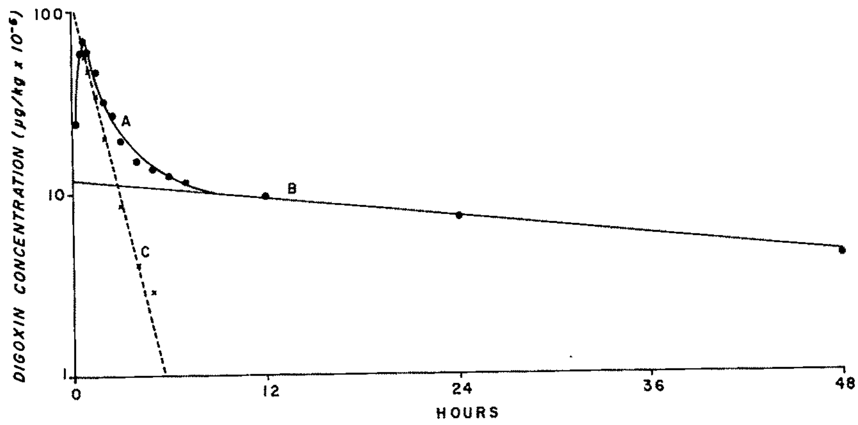


Figure 1 Composite (12 patients) serum concentration of digoxin following oral administration. Serum concentration is plotted on the vertical axis; time on the horizontal axis on a semilogarithmic scale. Although studies were continued for seven days, only the first two days are shown in order to demonstrate the important early differences in the serum turnover curves. Note that serum concentration reaches a peak 45 to 60 min after being administered. *Line B* is the best straight line that can be drawn back to zero time and represents the dominant T½ of digoxin (31.3 hr) and is associated with metabolism and excretion of the glycoside. *Line C* is derived by subtracting *line B* from the descending limb of *curve A,* thus eliminating metabolism and excretion. This exponential function now represents distribution and binding of the glycoside to the tissue and has a T½ of 60 min (reproduced with permission from Doherty, J. E., Perkins, W. H., Mitchell, G. D. 1961. *Arch. Intern. Med.* 108:351).

Two exponential functions may be derived from this disappearance curve. The first of these is attributed to the distribution and binding of the glycoside and has a half time (T½) of 60 min, while the second represents the dominant T½ of 31.3 hr and is attributed to the metabolism and excretion (primarily the latter) of tritium-labeled digoxin. There is obvious clinical correlation with the early onset of action of digoxin seen in the human subject. The dominant T½ also correlates with the disappearance of pharmacologic activity when the drug is withdrawn and is consistent with the term "short duration of action" frequently applied to this glycoside.

The excretion of digoxin is primarily in the urine as unchanged digoxin and is shown in Figure 4 along with excretion by other routes of administration.

DIGOXIN TURNOVER KINETICS AFTER INTRAMUSCULAR ADMINISTRATION

Tritiated digoxin was given to ten patients with congestive heart failure by the intramuscular route (10) and differences in the serum turnover curve are shown in Figure 2.

Intramuscular administration reveals a dominant T½ of 38 hr, a figure not statistically different from that obtained by oral administration. Note, however, that the peak serum concentration is later and somewhat lower than that found by the

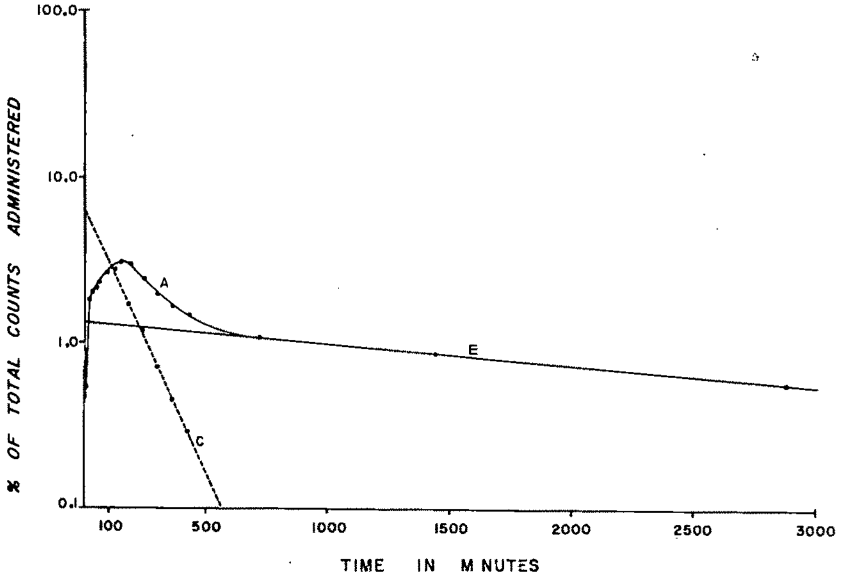


Figure 2 Composite (10 patients) serum concentration after i.m. administration. Plotted as described for Figure 1 with serum concentration on the vertical axis, time on the horizontal. *Lines B* and *C* are derived as on the preceding figure as well. Note that when digoxin is given by the i.m. route of administration, the peak serum level is later and lower than by the oral route of administration. The serum plateau is achieved only after about 10 to 12 hr, compared with 5 to 6 by the oral route. The dominant T½ (*line B*) is 38 hr and *line C* representing tissue distribution and binding is 100 min (reproduced with permission from Doherty, J. E., Perkins, W. H. 1965. *Am. J. Cardiol.* 15:170).

oral route. The serum plateau, which represents equilibration of the serum digoxin with tissue stores, is delayed to 10–12 hr. The early exponential (tissue distribution and binding phase) has a T½ of 100 min, nearly twice as long as by the oral route. This is probably related to two factors: (*a*) tritiated digoxin given orally was an alcoholic solution which may have facilitated absorption and thereby more rapid distribution to tissues, and (*b*) the gluteal muscle injection of digoxin in 40% propylene glycol, 10% alcohol, and 50% distilled water, besides being very painful, is not rapidly absorbed (probably because of this vehicle and possibly because of sequestration of digoxin at the injection site). Excretion of digoxin is again primarily in the urine and is shown in Figure 4.

DIGOXIN TURNOVER KINETICS AFTER INTRAVENOUS ADMINISTRATION

Thirteen patients were given tritiated digoxin intravenously, serum and excreta were collected as described previously, and the serum turnover curve for digoxin administered by this route is reproduced in Figure 3 (11).

There are three exponential functions which may be derived from this curve. An early curve representing serum distribution may be constructed having a T½ of 2

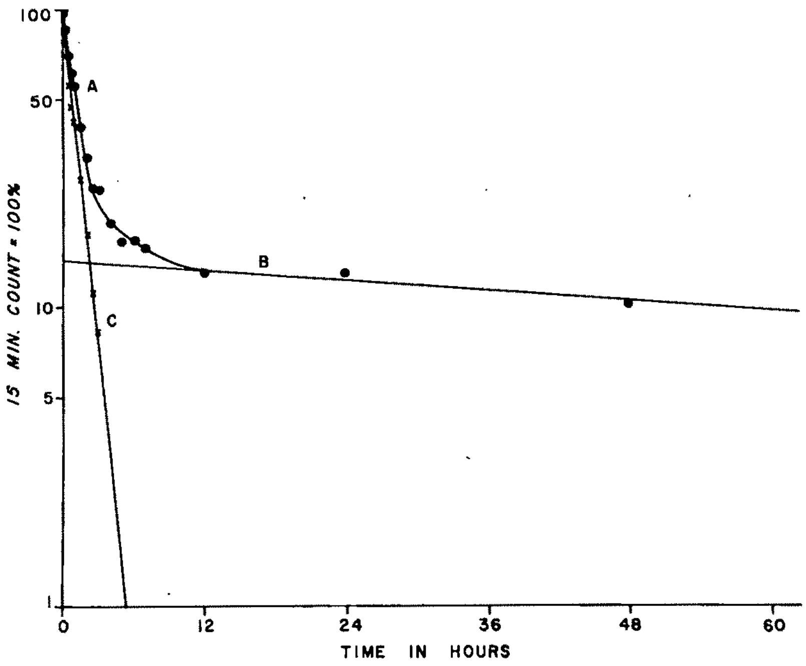


Figure 3 Composite (13 patients) serum concentration of digoxin after i.v. administration. The figure is plotted as are Figures 1 and 2. Note that high initial serum concentration follows i.v. use. A more rapid fall to the serum plateau is achieved than after oral or i.m. use (note that the scale is different). *Line B* represents metabolism and excretion of digoxin and T½ is 33.1 hr. *Line C* represents tissue distribution and binding and has a T½ of 30 min (only the first 60 hr of a 7-day study are shown). Reproduced with permission from Doherty, J. E., Perkins, W. H. 1962. *Am. Heart J.* 63:528.

min. It is not meaningful because of variations in injection time, blood volume, cardiac output, etc, and is not shown in the figure.

The organ distribution and binding T½ is shortened to 30 min, indicating the more rapid availability of digoxin by this route. The dominant T½ is 33.1 hr.

One may see that the dominant half time is not altered by the route of administration (the mean is 34 hr) but the early distribution is affected. These early differences are well shown in Figure 4. This demonstrates comparative serum concentrations obtained from oral, intramuscular, and intravenous routes of administration plotted as before and the 7-day excretion of digoxin in urine and stool.

Clinically important is the arrival at the serum plateau level: 5–6 hr following oral administration, 10–12 hr following intramuscular administration, and 2–4 hr following intravenous administration. Serum samples must be obtained at least this long after administration by the appropriate route for evaluation of clinical state of the patient with a digoxin serum assay (12). These times also represent maximum therapeutic activity after a given dose of digoxin by that route of administration.

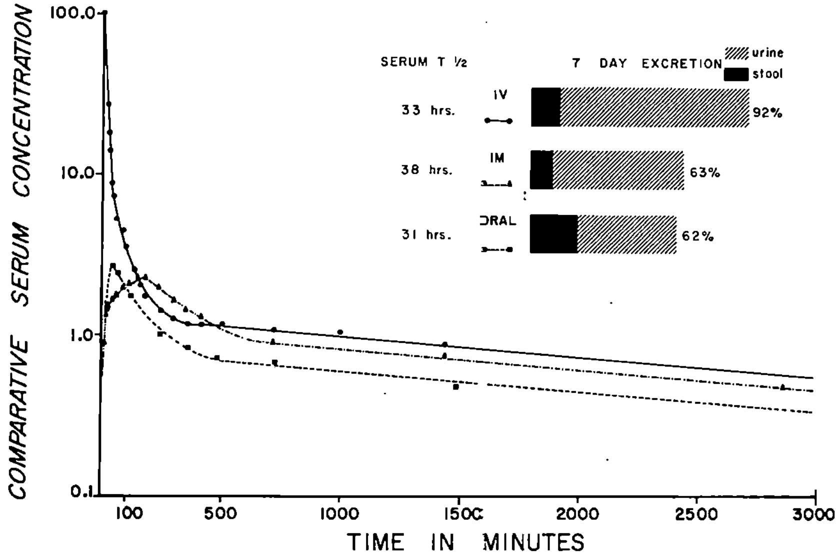


Figure 4 Comparative composite serum levels of digoxin and 7-day excretion rates after oral, i.m., and i.v. administration. The serum curves on the three preceding figures are shown as comparative serum concentration to illustrate the important early differences in serum concentration by the various routes of administration (see text). Excretion rates shown are for patients with normal renal function and clearly show the major route of excretion to be the kidney, regardless of the route of administration (reproduced with permission from Doherty, J. E. 1969. *J. Ark. Med. Soc.* 66:120–25).

ABSORPTION OF DIGOXIN The absorption of digoxin was calculated and found to be 85% in patients with normal renal function who are euthyroid and without malabsorption or electrolyte problems (13). It should be noted that this figure represents digoxin in its best biologically available form—in an alcoholic solution or dried on a sucrose cube. Digoxin tablets (Lanoxin®) are felt by the authors to be 60–75% absorbed.

Lindenbaum et al (14) noted variation in biologic availability of digoxin tablets in 1971. This led the U.S. Food and Drug Administration to adopt a tablet dissolution test (15) which all products must now adhere to and which insures adequate bioavailability on the American market.

EXCRETION OF DIGOXIN Refer again to Figure 4, which shows that excretion of digoxin by any route of administration is predominantly renal. Approximately 70% of an administered dose may be recovered from the urine in a 7-day study, 95–98% as unchanged digoxin. Ten to twenty percent digoxigenin bisdigitoxiside, 25% digoxigenin monodigitoxiside (13), and 45–55% digoxin may be recovered from the stool.

Digitoxin Pharmacokinetics

Digitoxin was the sixteenth most frequently prescribed drug in the U.S. in 1971 (1). The pharmacology of this valuable long acting glycoside has been well studied utilizing ^{14}C and ^{3}H labels by Okita (16, 17), double isotope dilution by Lukas (18), and inhibition red blood cell uptake of rubidium by Gjerdrum (19).

Virtually all of the kinetic studies with digitoxin reveal the long T½ which characterizes its clinical behavior in the human subject. Figure 5 from Lukas (18) summarizes this with great clarity. Half times of digitoxin are plotted in urine, stool, and serum in a human subject after discontinuing the drug after a single 1.0-mg dose (*left*) and after maintenance therapy of 0.1 mg daily (*right*). Note that the T½ as measured in excreta and that in sera is about 5 days, a figure corresponding to the half-life of the drug observed clinically as well as by serial changes in the ejection time index, a measure of inotropic activity. Because of the long dominant biologic half-life of digitoxin, differences which were noted with serum digoxin after various routes of administration are not important. It should be remembered that the onset of action of digitoxin is slightly prolonged when compared to the shorter acting glycosides (20).

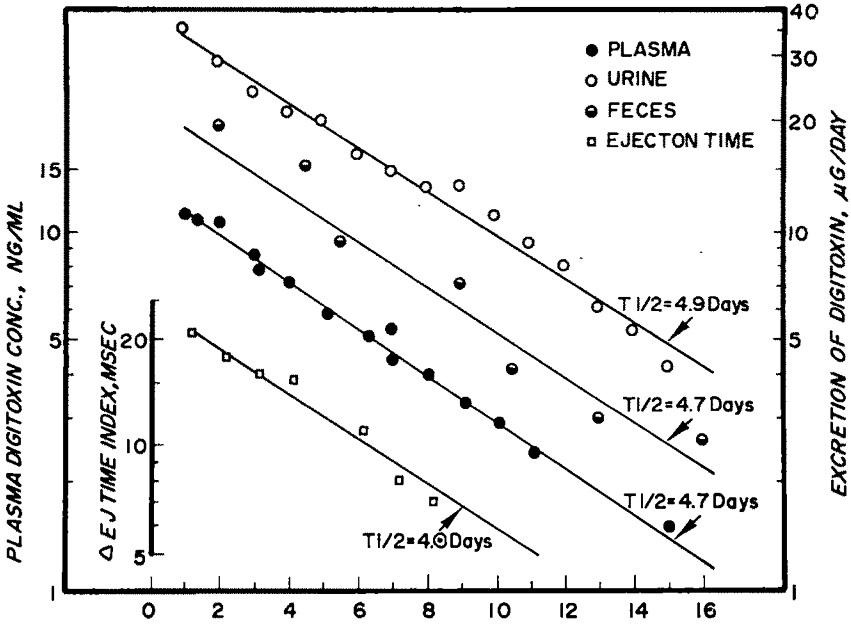


Figure 5 Plasma concentration and daily rates of urinary and fecal excretion of digitoxin and change in left ventricular ejection time index (ΔETI) on the vertical axis in a 65 year old man after 1.0 mg digitoxin, time in days shown horizontally. Reproduced with permission from D. S. Lukas (19).

It is prudent to obtain levels of serum digitoxin at least 6–8 hr after the last dose (orally or parenterally) in order to obtain a level that would be representative of myocardial digitoxin. This is, of course, not as important as with the "short-acting" glycosides.

ABSORPTION OF DIGITOXIN Digitoxin is presumed to be nearly completely absorbed (90–100%) as assessed by the effect of similar doses of the drug in slowing rate in patients with atrial fibrillation. Biologic availability has not been a problem both because of this and the long biologic T½ of digitoxin.

EXCRETION OF DIGITOXIN Lukas (18) estimated that approximately 14% of the total body pool of digitoxin is excreted daily, 4% in urine, 2% in feces, and 8% in biotransformation (to metabolites).

Other Cardioactive Glycosides

The two preparations discussed constitute 95% of digitalis glycosides used in the U.S. It should be remembered that digitalis leaf, which is still used by many physicians, is mostly digitoxin, but does contain many other substances as well. It has a long 5–7 day T½ and kinetic principles of digitoxin may be roughly applied.

Ouabain is a shorter duration of action glycoside than digoxin, its dominant T½ being 21 hr and its excretion primarily renal (21). It has a more rapid onset of action and achieves peak inotropic activity faster than digoxin, deslanoside, and digitoxin (20).

Deslanoside, a popular glycoside for parenteral use, has a T½ about that of digoxin (22), and is excreted primarily in the urine. Its use in maintenance therapy is not recommended.

SPECIAL CLINICAL PROBLEMS WITH DIGITALIS THERAPY

Renal Disease

The presence of renal insufficiency and an elevation of the blood urea nitrogen beyond 50 mg% (provided prerenal azotemia is excluded) or a reduction in the creatinine clearance indicates a compromise in digoxin excretion (23). A downward adjustment in dosage is necessary to avoid toxicity. This problem is not as great with digitoxin. Storstein (24) has demonstrated in renal failure that more digitoxin is metabolized to digoxin whose T½ in the anephric state approximates that of digitoxin with normal renal function. We cannot predict the individual rate of metabolic conversion, and the state of digitalization achieved with the two glycosides "on board" might be difficult to assess.

Patients subjected to peritoneal or hemodialysis on digitalis do not lose appreciable amounts of the drug by this route and these techniques should not be used in the management of digitalis intoxication. It should also be remembered that patients maintained on chronic dialysis therapy required reduced doses of digoxin.

Electrolyte Disturbances

The increasing utilization of powerful diuretic agents in patients receiving digitalis often induces excessive potassium loss, thus sensitizing the myocardium to digitalis. Supplementary potassium may be necessary or perhaps the addition of a "potassium sparing" diuretic such as spironolactone or triamterene will suffice. Magnesium loss may also occur and if present, magnesium replacement should be accomplished. Calcium, on the other hand, sensitizes the myocardium to digitalis and digitalis should be used with caution in hypercalcemic states. Administration of calcium salts (particularly parenterally) is relatively contraindicated in digitalized patients.

Thyroid Disease

Patients with hyperthyroidism require proportionately more digoxin to achieve therapeutic effect because of lower serum levels of the drug. Patients with myxedema achieve higher serum digoxin levels for comparable doses and are sensitive to the drug. Should digitalis therapy be required in thyroid disease, increased digoxin doses are suggested in thyrotoxicosis, and reduced doses in hypothyroidism (25).

Liver Disease

Patients with hepatic insufficiency have been shown to metabolize and excrete digoxin in a normal fashion, indicating the presence of significant extrahepatic sites of metabolism as well as demonstrating the lack of digoxin metabolism as a determinant of its half-life (26).

Digitoxin, having a more important metabolic role involved in its excretion, would be expected to experience less biotransformation and perhaps accumulation of cardioactive metabolites. Smaller doses of digitoxin are suggested in hepatic insufficiency (27).

Malabsorption

Malabsorption of digoxin has been reported to occur with patients having sprue (28) but has recently been shown not to be a very important factor in patients with other rapid transit syndromes (28). Adjustments in dose level should be monitored by serum assay and careful clinical assessment.

Elderly Patients

Ewy et al (30) demonstrated that elderly patients' requirements for digoxin were reduced because of a decrease in lean body mass (skeletal muscle is the major depot of digoxin) and reduction in creatinine clearance. Smaller doses are recommended in the elderly.

Obese Patients

Similarly, little digoxin is present in adipose tissue, so lean body mass should be used to estimate digoxin requirements when digitalization is required in the obese patient (31).

DRUG INTERACTIONS

The diuretic agents previously mentioned as responsible for hypokalemia, and possibly hypomagnesemia (both drug-patient-drug) are important interactions. Cholestyramine and colestipol bind the glycosides in the gastrointestinal tract and prevent absorption (either primary or enterohepatic). Neomycin appears to limit absorption of digoxin to a modest degree (32). Microsomal enzyme induction of digitoxin metabolism to digoxin may result in increased dosage requirements in patients receiving diphenylhydantoin, phenobarbital, and phenylbutazone in addition to digitoxin (33). This interaction results in an increased digitoxin requirement because of the much shorter T½ of digoxin. Propranolol and reserpine in combination with digitalis may induce excessive bradycardia.

SERUM LEVELS OF DIGITALIS

The demonstration of a reasonable relationship between serum levels and myocardial levels of digoxin (25) suggested that if practical methods of serum assay could be devised, clinical benefit would result. Several serum assays are now available (3). The most useful of these appears to be the radioimmunoassay developed by Smith, Butler & Haber (34). Table 1 outlines appropriate serum levels of digoxin and digitoxin for clinical application. Note that there is some overlap between normal and abnormal levels. One should remember that "increased" serum levels of digoxin are sometimes seen without toxicity in patients with atrial arrhythmias or hyperkalemia, and in infants and children. Normal or low serum levels *with* toxicity are often seen with hypokalemia, hypomagnesemia, lung scans (or isotopes "on board"), myocardial infarction, hypoxia, and myxedema.

Table 1 Therapeutic and toxic serum levels of digitalis glycosides

	Therapeutic (Overlap) ng/ml		Definitely Toxic
Digitoxin	15 - 30	30 - 45	45.0+
Digoxin	0.5 - 2.5	2.0 - 3.0	3.0+

Low serum levels are seen with large oral doses in patients with hyperthyroidism, intestinal malabsorption, and abnormal metabolism.

Perhaps the most appropriate use of the digitalis serum level is not in the patient who is toxic, but in the patient who is underdigitalized. An increased daily dose of the drug results in improved control of this patient and the dose tailored to the serum level and the clinical response.

DIGITALIS INTOXICATION

Digitalis toxicity is one of the most common and potentially dangerous adverse drug reactions. The incidence of toxicity is estimated to be between 15 and 20% (35).

Recent studies from various institutions have suggested a mortality rate from digitalis toxicity between 7 and 50%. Beller et al (36) indicate that the mortality in a group of digitalis-intoxicated patients is roughly twice that of similar patients on digitalis who were not intoxicated.

There is still some speculation about the mechanism of digitalis toxicity. Digitalis causes an efflux of potassium from the myocardial cell (37). This is attributed to the inhibitory effect of the cardiac glycosides on membrane sodium-potassium ATPase. Inhibition of this enzyme causes a loss of intracellular potassium and a decrease in the resting membrane potential. This results in a slower rate of rise of the action potential, a slower conduction velocity (38), and increased spontaneous depolarization. This mechanism can explain the re-entrant tachyarrhythmias, the enhanced ventricular ectopic activity, and depression of A-V conduction that is seen with toxic concentrations of digitalis.

Recognition of Digitalis Toxicity

Widespread use of more purified digitalis preparations has made gastrointestinal and other nonspecific symptoms of toxicity less helpful, and often cardiac arrhythmia is the only clue to diagnosis. About 80% of the intoxicated patients may have arrhythmia and conduction disturbances. The arrhythmia usually is a combined effect of the drug on impulse formation and conduction. While almost any arrhythmia can result, it is useful to note that there are several abnormal rhythms seldom seen in digitalis intoxication. Mobitz II A-V block, parasystolic rhythm, chaotic atrial tachycardia, and atrial flutter and fibrillation with a rapid ventricular response are distinctly uncommon (39).

The incidence of ventricular arrhythmias associated with digitalis intoxication exceeds all other rhythm disturbances. Premature ventricular contractions (PVC) are most common, and at the same time the least specific (40). When PVC occur in a bigeminal or multifocal pattern, digitalis may be responsible, but at the same time they may be due to advanced cardiac disease. A digitalis serum level is useful in this difficult problem (Table 1). It is helpful if the PVC are associated with conduction defects or other ectopic rhythms. Chung (41) reported 180 cases of digitalis toxicity with an overall incidence of ventricular arrhythmia of 57.8%. Ventricular tachycardia, defined as three or more consecutive PVC, was seen in 9%. Bidirectional ventricular tachycardia, which may represent alternating pathways of conduction or alternating ectopic impulses, is reasonably specific for digitalis intoxication. Any of these may terminate in ventricular fibrillation and sudden death.

Atrial tachycardia with block is a somewhat uncommon digitalis-induced arrhythmia, but is rather specific and usually associated with a poor prognosis. Nonparoxysmal junctional tachycardia, however, is a more common arrhythmia due to digitalis. This arrhythmia is often unrecognized, particularly when associated with aberrant intraventricular conduction. Differentiation of nonparoxysmal junctional tachycardia with aberrant conduction and ventricular tachycardia may be difficult or impossible. The rhythm is generally absolutely regular, but in the presence of exit block may resemble a Mobitz II or Wenckebach phenomenon. The presence of group beating should alert one to the possibility of a junctional tachycardia with exit block. Nonparoxysmal junctional tachycardia may also co-exist with sinus rhythm, atrial tachycardia, or atrial flutter. Chung found 87 patients with atrial fibrillation who were toxic to digitalis and 83 of these showed either nonparoxysmal junctional tachycardia or junctional escape rhythm (41). The presence of a regular ventricular response with atrial fibrillation on the electrocardiogram should be interpreted as evidence of digitalis intoxication until proven otherwise.

Disturbances of sinus impulse formation and conduction are much less common than ventricular arrhythmias in digitalis toxicity. The significance of first degree A-V block as a criteria for digitalis excess may be questioned as P-R prolongation may represent digitalis effect, but not necessarily toxicity unless a more serious rhythm disturbance is present. A-V conduction disturbance is fairly common with second degree A-V block of the Wenckebach-type being most frequently observed. Complete A-V block is uncommon unless there is pre-existing atrial fibrillation.

Treatment of Digitalis Toxicity

A specific discussion of treatment of digitalis toxicity is beyond the scope of this review. Suffice it to say many arrhythmias can be managed effectively by simply discontinuing the drug. Pharmacologic management may be necessary for more serious problems and some may require temporary pacemakers for heart block.

ACKNOWLEDGMENT

We wish to acknowledge the secretarial assistance of Ms. Faye Hill, the technical assistance of Ms. Jacquelyn Gammill and Ms. Joyce Sherwood, encouragement from Drs. Robert Abernathy, Joseph Bates, and Marvin Murphy, and inspiration from our housestaff and students at the University of Arkansas.

Literature Cited

1. 1974. *National Prescription Audit,* p. 27. Ambler, Pa.: Gosselin. 9th ed.
2. Moe, G. K., Farah, A. E. 1970. *The Pharmacological Basis of Therapeutics,* ed. L. S. Goodman, A. Gilman, 677–708. New York: MacMillan. 4th ed.
3. Smith, T. W., Haver, E. 1973. *N. Engl. J. Med.* 289:945–52, 1010–15, 1063–72, 1125–29
4. Skou, J. C. 1972. *Basic and Clinical Pharmacology of Digitalis,* 48–69. Springfield, Ill.: Thomas
5. Okita, G. T., Richardson, F., Roth-Schechter, B. F. 1973. *J. Pharmacol. Exp. Ther.* 185:1–11
6. Przybyla, A. C., Paulay, K. L., Stein, E., Damato, A. N. 1974. *Am. J. Cardiol.* 33:344–50
7. Lowenstein, J. M., Corrill, E. M. 1966. *J. Lab. Clin. Med.* 67:1048–52
8. Smith, T. W. 1970. *J. Pharmacol. Exp. Ther.* 175:352–59
9. Doherty, J. E., Perkins, W. H., and Mitchell, G. D. 1961. *Arch. Intern. Med.* 108:531–39
10. Doherty, J. E., Perkins, W. H. 1965. *Am. J. Cardiol.* 15:170–74
11. Doherty, J. E., Perkins, W. H. 1962. *Am. Heart J.* 63:528–36
12. Doherty, J. E., Hall, W. H., Murphy, M. L., Beard, O. W. 1971. *Chest* 59:433–37
13. Doherty, J. E. et al 1970. *Circulation* 42:867–73
14. Lindenbaum, J., Mellow, M. H., Blackstone, M. O., Butler, V. P. Jr. 1971. *N. Engl. J. Med.* 285:1344–47
15. Lindenbaum, J., Butler, J. P. Jr., Murphy, J. E., Cresswell, R. M. 1973. *Lancet* 1:1215–19
16. Okita, G. T., Talso, P. J., Curry, J. H., Smith, F. D. Jr., Geiling, E. M. K. 1955. *J. Pharmacol. Exp. Ther.* 113:376–82
17. Okita, G. T., Talso, P. J., Curry, J. H. Jr., Smith, F. D. Jr., Geiling, E. M. K.

1955. *J. Pharmacol. Exp. Ther.* 115: 371–79

18. Lukas, D. S. 1973. *Digitalis,* ed. O. Storstein, 84–102. Oslo, Norway: Gyldendal Norsk Forlag
19. Gjerdrum, K. 1973. See Ref. 18, pp. 103–12
20. Forester, W., Lewis, R. P., Weissler, A. M., Wilke, T. A. 1974. *Circulation* 49:517–21
21. Seldon, R., Smith, T. W. 1972. *Circulation* 45:1176–82
22. Weissler, A. M., Snyder, J. R., Schoenfeld, C. D., Cohen, S. 1966. *Am. J. Cardiol.* 17:768–80
23. Doherty, J. E., Perkins, W. H., Wilson, M. C. 1968. *Am. J. Med.* 37:536–44
24. Storstein, L. 1973. See Ref. 18, pp. 158–68.
25. Doherty, J. E., Perkins, W. H. 1966. *Ann. Intern. Med.* 64:489–507
26. Marcus, F. I., Kapadia, G. G. 1964. *Gastroenterology* 47:517–24
27. Lukas, D. S. 1973. See Ref. 18, pp. 192–99
28. Heizer, W. D., Smith, T. W., Goldfinger, S. E. 1971. *N. Engl. J. Med.* 285:257–59
29. Hall, W. H., Doherty, J. E. 1974. *Am. J. Med.* 56:437–42
30. Ewy, G. A., Kapadia, G. G., Yao, L., Lullin, M., Marcus, F. I. 1969. *Circulation* 39:449–53
31. Ewy, G. A. et al 1971. *Circulation* 44:810–14
32. Bigger, J. T. Jr., Strauss, H. C. 1972. *Semin. Drug Treat.* 2:147–77
33. Solomon, H., Abrams, W. B. 1972. *Am. Heart J.* 83:277–80
34. Smith, T. W., Butler, V. P. Jr., Haber, E., 1969. *N. Engl. J. Med.* 281:1212–16
35. Mason, D. T. et al 1971. *Am. J. Cardiol.* 27:546–59
36. Beller, G. A., Smith, T. W., Abelmann, W. H., Haber, E., Hood, W. B. Jr. 1971. *New Engl. J. Med.* 18:989–97
37. Roberts, J., Kelliher, G. J. 1972. *Semin. Drug Treat.* 2:214–19
38. Dreifus, L. S., Watanabe, Y. 1972. *Semin Drug Treat.* 2:194–98
39. Fisch, C., Knoebel, S. B. 1970. *Progr. Cardiov. Dis.* 13:71–96
40. Chung, E. K. 1970. *Am. Heart J.* 79:845–48
41. Chung, E. K. 1969. *Digitalis Intoxication.* Baltimore: Williams & Wilkins

Copyright 1975. All rights reserved

CALCIUM OXALATE RENAL STONES

❖7140

Edwin L. Prien, Jr., M.D.
Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

Modern Western stone disease is composed of four major stone entities, with the majority of patients ($\sim$ 70%) suffering from what has been loosely defined as "calcium stone disease" (1). For this reason it is especially distressing that the pathogenesis and therapy of this disorder has remained so elusive. There can be little doubt that cystine, uric acid (not urate), and mixed magnesium ammonium phosphate/apatite ("infected") stones are a consequence of excessive supersaturation of the urine for the offending species (2). Such clear-cut excessive supersaturation is not always demonstrable in calcium stone disease although it is generally suspected when either marked hypercalciuria or primary hyperoxaluria are present. In this short review I will attempt to deal with this problem and comment on the current status of the older theories of pathogenesis as well as several newer concepts.

All theories of stone formation require a certain degree of urine supersaturation if not for spontaneous nucleation at least for continued stone growth. In order to determine the degree of supersaturation, a specific solid phase (crystal) must be anticipated. Calcium oxalate (mono- and dihydrate) and apatite (a complex calcium phosphate similar to bone salt) comprise at least 90% of calcium stones. They are either pure calcium oxalate or calcium oxalate with variable amounts of apatite intermixed. This apatite may be nuclear, or may appear in scattered nests or in layers alternating with calcium oxalate. Approximately 5% are pure apatite and 3% contain brushite ($CaHPO_4 \cdot 2H_2O$) either solely or admixed with other constituents (1).

URINARY CALCIUM

Whatever the variety it is apparent that calcium is common to all these crystalline compounds. This, together with the relative ease of calcium measurement and the observation that stones are sometimes associated with disorders of calcium metabolism has led to the assumption that all calcium stones are the result of abnormalities of calcium metabolism. It can safely be said that many stone formers have an elevation of calcium excretion and presumably of urinary calcium concentration. A chief difficulty in the refinement of this notion has been in defining the normal range. If it is taken as the range of excretion for 90% of the population, then a reasonable upper limit of 300 mg per day for men and 275 mg for women has been suggested. However, more recent data indicate levels as high as 400 mg for men (3). If a different definition of normal is used, namely the absence of stone, then clearly there are individuals with still higher urinary calcium who do not suffer stone disease. Are such people "normal" or should they be considered "stone formers" who lack some other factor necessary for stone production? It would also be important to be sure they are not relatives of stone formers, a simple precaution generally overlooked.

A second difficulty in assessing urinary calcium excretion, apart from its intrinsic daily variation, is its relation to dietary calcium. In the majority of normals urine calcium is little affected by dietary calcium. However, in perhaps a majority of stone formers there is a relatively steep rise (slope) of calcium excretion with increasing calcium intake. Thus on low calcium diets most stone formers have normal urinary calcium while on high intakes the majority are elevated (3). No doubt simple dietary restriction accounts for the multitudes of young men who have but one stone per lifetime.

The riddle is why stone disease continues to recur in some patients whose urine calcium has been normalized on low calcium diets or in patients who have always had low normal urine calciums (100–200 mg per day) not to mention why some remain stone-free with much higher calciums This dilemma has led investigators to search for other factors.

INHIBITORS OF CRYSTALLIZATION

Normal urine has an abundance of inhibitors of crystal formation including ionic strength, various complexers, and certain "crystal poisons" that function in exceedingly low concentrations (10^{-6} M). Studies utilizing the rachitic rat cartilage assay suggested that a substance, possibly a peptide was deficient in stone formers and the concept of "evil urine" was born (4). To date no specific deficient inhibitor has been isolated. One inhibitor that has been well characterized is inorganic pyrophosphate, but it is not deficient. Recently, the specific aspects of calcium oxalate crystal growth and aggregation (not nucleation) have been studied (5). Again stone-forming urine appears to show a deficiency of inhibition, but whether this effect is due to lack of an inhibitor or whether the more rapid rates of crystal growth observed are a consequence of higher degrees of supersaturation remains to be seen.

STONE MATRIX

An alternative theory has evolved from the observation that all urinary calculi have a mucoprotein matrix accounting for ~ 3% of stone weight (6). It is possible that matrix might function as a specific template (by epitaxy) for nucleation of crystalline components from normally supersaturated urines, although its occurrence in stones of all compositions is an argument against specificity. Initially, the major component of matrix, characterized immunologically as matrix substance A, was found only in the urine of stone formers and not normals. It has since been found in other nonstone kidney diseases and appears to be related to renal inflammation or destruction rather than to stone (6).

URINE SUPERSATURATION

In studies which can only be described as elegant and comprehensive the Leeds group has made important recent observations. By counting the number and volume of calcium oxalate crystals in fresh warm urine (37°C), only stone formers are found to have large and aggregated crystals, while both they and normals may have much smaller crystals averaging 5 μ in size (7). They then undertook the complex calculation (computer-assisted) of the degree of supersaturation (ion products) for calcium oxalate in these urine specimens. The greater crystal volumes in stone formers did correlate with greater peak supersaturation. This result was due simply to greater concentrations of calcium and oxalate and not to differences in other ion species, complexers, or ionic strength (8).

URINARY OXALATE

The Leeds group then made what may prove to be a seminal observation. With the computer they compared the effect on supersaturation of either increasing calcium concentration or increasing oxalate concentration. Surprisingly, increases of oxalate were more potent in raising supersaturation while additions of calcium were largely complexed (3). Here may be an eloquent answer to the riddle of why stone disease correlates so incompletely with urinary calcium. It remains only to show that stone formers have higher 24-hr urinary oxalate excretions than do normals. However, this has been conclusively shown only for primary hyperoxaluria and for the newer entity of hyperoxaluria secondary to disease (or resection) of the terminal ileum (9). With one exception, numerous investigators including the Leeds group have found at most only slightly higher mean values in stone formers with a wide range of overlap. However, a group in Paris has performed over 20,000 such determinations using gas-liquid chromatography and claims elevations in the majority of patients (10).

This problem is potentially one of great importance and subtlety. Not only is oxalate notoriously difficult to measure, but it is present in low concentrations (one tenth that of calcium on a molar basis), and small differences may be significant. Further, the daily variation is considerable even on a constant diet, and perhaps most important it is a function of dietary calcium intake. Urinary oxalate excretion can be shown to rise with calcium restriction (11). This phenomenon would provide a tidy explanation of stone recurrence in a patient whose initial stone was due mainly to hypercalciuria and whose subsequent stone on a low calcium diet might be due to secondary elevations of urinary oxalate. While the potential role of oxalate in stone disease is intriguing, there is little current data based on 24-hr urine collections to support it.

POSSIBLE ROLE OF CALCIUM PHOSPHATE

So far nothing has been said of the calcium phosphate (apatite) that occurs either with calcium oxalate or as pure apatite stones. A feature common to all urinary phosphates is decreasing solubility with increasing pH. The evidence suggests that the stones in distal renal tubular acidosis are essentially pure apatite. The persistence of relatively more alkaline urine than normal (pH > 6.0) is no doubt fundamental, although it is clearly not corrected with alkalinization therapy directed at the systemic acidosis and related hypercalciuria and low citrate excretion.

However, since renal tubular acidosis is a rare disorder, the majority of pure or predominantly apatite stones are found in the large group of calcium stone formers. These patients appear to have generally more alkaline urines than do their counterparts with pure or predominantly calcium oxalate stones (12). It is not clear that gentle measures to acidify their urines would do anything more than change their stone composition to purer calcium oxalate, as the Leeds group finds as high peak urine calcium and oxalate concentrations as in the other stone formers (12). This admixed calcium phosphate may simply represent crystalluria which has been caught up in a calcium oxalate stone during alkaline tides.

In this context an alternative theory has been proposed (13). Since at least 50% of calcium oxalate stones contain variable amounts of apatite, occasionally in a "nuclear" location, the suggestion has been made that it may well be this calcium phosphate that nucleates and "governs" the development of calcium oxalate calculi. By a mechanism analogous to one theory of bone formation, it is postulated that brushite ($CaHPO_4$) is the antecedant of the apatite in calculi and that supersaturation of the urine for brushite may be the fundamental abnormality in calcium oxalate stone formers. The brushite would nucleate the calcium oxalate and then transform into apatite by hydrolysis. This transformation in urine is generally accepted and would account for the rarity of brushite in stones. It can also be shown that brushite will nucleate calcium oxalate.

However, it is the application of this theory to patients and to urine that remains at issue. A potent regulator of brushite supersaturation is the pH of that urine. In the absence of hypercalciuria, urine is excessively supersaturated with brushite only when relatively alkaline. This is true of both stone formers and normals (13). In the Leeds study more alkaline urine is associated with the appearance of apatite in the calcium oxalate stones (12). So while such supersaturation for brushite accounts nicely for the apatite in stones, the dilemma is to show that it has anything at all to do with the calcium oxalate, especially in patients with normally acid urines. Finally, it seems likely that it is not the *nucleation* of calcium oxalate that is critical, as normal individuals have calcium phosphate crystalluria during alkaline tides as well as the small calcium oxalate crystals noted above. What is important for stone production are factors that promote stone *growth* of the relevant species, namely, calcium oxalate. These factors must be either calcium, oxalate, or lack of growth inhibitors.

POSSIBLE ROLE OF URIC ACID

Another theory that represents a unique departure derives from the clinical observation that hyperuricemia is associated with calcium stone disease (14). In fact, given that calcium stones are about twelve times more common than uric acid stones and that only a minority of patients with uric acid stones have hyperuricemia, the likelihood is that a hyperuricemic stone patient will be a calcium oxalate stone former. This phenomenon has created the uneasy feeling that somehow uric acid may be nucleating calcium oxalate. However, it can be categorically stated that such stones do not have uric acid nuclei. The association of these two components is uncommon. A recent study has shown that some calcium stone formers do have small elevations of uric acid excretion even in the absence of hyperuricemia (15). But whether this is simply a consequence of diet or altered uric acid metabolism or renal handling is unclear, and whether it contributes to calculus production is simply speculation. If the notion is that exceedingly small uric acid crystals are nucleating calcium oxalate, then data concerning the possibility of more acid urines than normal should also be developed.

It is interesting to note that both these new theories of stone pathogenesis (brushite theory, uric acid theory) have relegated calcium oxalate to the role of the innocent bystander. In effect this is only possible by placing constraints on urinary pH. A tendency to more alkaline urine in the former and, by implication, more acid urine in the latter, is required. For calcium oxalate stones the dilemma is twofold: on the one hand it will be necessary to show that these phenomena are present without simply resulting in the appearance of significant amounts of apatite or uric acid in calcium oxalate stones, and on the other hand it will remain difficult to show that vanishingly small amounts of either apatite or uric acid control the growth of a stone that is greater than 95% calcium oxalate.

TREATMENT

In the treatment of recurrent calcium stone disease several facts emerge. There are a host of therapeutic regimens and their theoretical utility depends on widely differing mechanisms. Yet, their proponents all claim effectiveness. In a disorder known for its intermittency and unpredicability, it is unfortunate that none of the therapeutic trials has a control group. The only control has been the patient's stone history prior to therapy. Also it is only human nature that patients will embark on therapy after having just passed one or more stones. One inescapable interpretation is that a control group would do just as well when being followed in a program by enthusiastic clinicians while maintaining dietary restriction and forcing fluids. In fact this possibility would make the detection of a truely effective medicine in a properly controlled setting all the more difficult.

The first step in treatment is dietary calcium restriction and promotion of fluid intake. If stones continue to recur in spite of reductions of urinary calcium, then dietary oxalate restriction may be useful. Although such restriction has been shown to curtail secondary urinary oxalate elevations, adherence to such a regimen is difficult on an outpatient basis (11).

For those patients who fail on dietary therapy, any of the following programs has been recommended. Cellulose phosphate is effective in reducing urinary calcium but has not been released by the FDA (16). It does not prevent secondary oxalate elevations (17). Thiazides are well tolerated and reduce urinary calcium by 50% (18) and do not appear to be associated with rises in oxalate (17). Oral phosphate therapy reduces urine calcium only modestly while greatly augmenting urine phosphate and increasing urine oxalate (19). Any beneficial effect of this drug probably resides in increased urinary pyrophosphate excretion. Magnesium oxide has been advocated by virtue of its ability to prevent calculi in pyridoxine-deficient rats, although the mechanism is uncertain (20), and pyridoxine has been used in conjunction with magnesium in the hope of lowering urinary oxalate excretion, although this effect has been inconstant (21). Allopurinol has been claimed effective although no entirely logical mechanism is evident (22). It might be related to decreases in urinary uric acid. Interestingly, xanthine oxidase is one of the three enzymes involved in the conversion of glyoxalate to oxalate. But this pathway appears to be of minor importance and the administration of allopurinol does not reduce urinary oxalate (23). Finally, succinimide (not available in the USA) lowers urinary oxalate perhaps by condensing with it and allowing it to be consumed by the Krebs cycle (10).

Clearly, there is much work still to be done on the mechanisms underlying recurrence of calcium oxalate stones as well as its proper therapy. The question is of sufficient subtlety that it is unlikely that animal models will be useful as the one animal which most closely resembles the stone former is the normal human nonstone former.

Literature Cited

1. Prien, E. L., Prien, E. L. Jr. 1968. Composition and structure of urinary stone. *Am. J. Med.* 45:654–72
2. Smith, L. H. Jr., Williams, H. E. 1971. Kidney stones. In *Diseases of the Kidney*, ed. M. B. Strauss, L. G. Welt. 973–96. Boston: Little, Brown. 2nd ed. 1456 pp.
3. Nordin, B. E. C. 1973. *Metabolic Bone and Stone Disease.* Baltimore: Williams and Wilkins. 309 pp.
4. Howard, J. E., Thomas, W. C. Jr. 1968. Control of crystallization in urine. *Am. J. Med.* 45:693–99
5. Robertson, W. G., Peacock, M. 1972. Calcium oxalate crystalluria and inhibitors of crystallization in recurrent renal stone formers. *Clin. Sci.* 43:499–506
6. Boyce, W. H. 1968. Organic matrix of human urinary concretions. *Am. J. Med.* 45:673–83
7. Robertson, W. G., Peacock, M., Nordin, B. E. C. 1969. Calcium crystalluria in recurrent renal stone formers. *Lancet* ii: 21–24
8. Robertson, W. G., Peacock, M., Nordin, B. E. C. 1971. Calcium oxalate crystalluria and urine saturation in recurrent renal stone formers. *Clin. Sci.* 40:365–74
9. Chadwick, V. S., Modha, K., Dowling, R. H. 1973. Mechanism for hyperoxaluria in patients with ileal dysfunction. *N. Engl. J. Med.* 289: 172–76
10. Thomas, J. et al 1973. The role of oxalic acid in oxalic nephrolithiasis. In *Proceedings of the International Symposium on Renal Stone Research,* ed. L. Cifuentes Delatte, A. Rapado, A. Hodgkinson. 57–66. Madrid, Spain: S. Karger. 363 pp.
11. Marshall, R. W., Cochran, M., Hodgkinson, A. 1972. Relationships between calcium and oxalic acid intake in the diet and their excretion in the urine of normal and renal-stone-forming subjects. *Clin. Sci.* 43:91–99
12. Marshall, R. W. et al 1972. The relation between the concentration of calcium salts in the urine and renal stone composition in patients with calcium-con-

taining renal stones. *Clin. Sci.* 43: 433–41

13. Pak, C. Y. C. 1969. Physiochemical basis for the formation of renal stones of calcium phosphate origin: Calculation of the degree of saturation of the urine with respect to brushite. *J. Clin. Invest.* 48:1914–22
14. Smith, M. J. V., Hunt, L. D., King, J. S. Jr., Boyce, W. H. 1969. Uricemia and urolithiasis. *J. Urol.* 101:637–42
15. Kavalich, A., Moran, E., Coe, F. 1973. Mechanism of hyperuricosuria in calcium stone formers. *Proc. Am. Soc. Nephrol.* 6:57 (Abstr.)
16. Pak, C. Y. C., Delea, C. S., Bartter, F. C. 1974. Successful treatment of recurrent nephrolithiasis (calcium stones) with cellulose phosphate. *N. Engl. J. Med.* 290:175–80
17. Marshall, R. W., Barry, H. 1973. Urine saturation and the formation of calcium-containing renal calculi: The effects of various forms of therapy. See Ref. 10, pp. 164–69
18. Yendt, E. R., Guay, G. F., Garcia, D. A. 1970. The use of thiazides in the prevention of renal calculi. *Can. Med. Assoc. J.* 102:614–20
19. Smith, L. H., Thomas, W. C. Jr., Arnaud, C. D. 1973. Orthophosphate therapy in calcium renal lithiasis. See Ref. 10, pp. 188–97
20. Melnick, I., Landes, R. R., Hoffman, A. A., Burch, J. F. 1971. Magnesium therapy for recurring calcium oxalate urinary calculi. *J. Urol.* 105:119–22
21. Prien, E. L., Gershoff, S. F. 1974. Magnesium oxide—pyridoxine therapy for recurrent calcium oxalate calculi. *J. Urol.* 112:509–12
22. Coe, F. L., Raisen, L. 1973. Allopurinol treatment of uric acid disorders in calcium stone formers. *Lancet* i:129–31
23. Gibbs, D. A., Watts, R. W. E. 1966. An investigation of the possible role of xanthine oxidase in the oxidation of glyoxalate to oxalate *Clin. Sci.* 31:285–97

Copyright 1975. All rights reserved

ALTERATIONS IN T- AND B-CELLS IN HUMAN DISEASE STATES[1]

❖7141

Ralph C. Williams, Jr., M.D. and Ronald P. Messner, M.D.
Department of Internal Medicine, University of New Mexico School of Medicine,
Bernalillo County Medical Center, Albuquerque, New Mexico 87131

INTRODUCTION

The emphasis in modern clinical immunology has recently shifted toward attempts to relate the immune response in various disorders to basic mechanisms of cellular immunology. In this context recognition of two major lymphocyte cell types, T- and B-cells, has served as a convenient focal point for examination of the immune system.

Lymphocytes have been divided into two major categories on the basis of physical and functional differences. B-cells are felt to represent precursors of antibody-forming cells. They are probably derived from or originate in the bone marrow and are characterized by the presence of surface immunoglobulins (1–7) which appear to function as antibody receptors for antigens. B-cells have also been identified by the presence of a receptor for the Fc fragment of antibody molecules complexed to antigens (8), and for the third component of complement (9, 10). Another method for B-cell identification utilizes binding of fluorescent aggregates of IgG to an aggregate receptor (11), presumably similar to the Fc receptor. The distribution of surface Ig, Fc receptors or aggregate receptors, and C3 receptors on B-cells is still not entirely clear. Recently some evidence has been presented that B-cell subpopulations may exist in which only a part of all the surface receptors are present (12) in that some B-cells may possess both Fc and C3 receptors whereas others show only C3 receptors. Since B-cells have been directly related to humoral antibody formation, the distribution of various B-cell receptors has also been proposed as a way of classifying various types of immune deficiencies (13).

T-cells are primarily influenced by the thymus. They bear no readily detectable amounts of surface immunoglobulin though a controversy exists over whether T-cell specificity is ultimately controlled by minute amounts of intrinsic or bound surface

---

[1] Supported in part by National Institutes of Health Grants AMAI 13824–04 and TO1 AI·00393–04A1A1D.

immunoglobulin (14–17). T-cells are primarily involved in various effector mechanisms of cell-mediated immunity (18–21) such as homograft rejection, mediation of delayed type hypersensitivity, or tumor cell destruction. T-cells can be identified using heterologous anti-T-cell antisera (22) or by formation of rosettes using sheep erythrocytes (23–25). The scanning electron microscopic appearance of T-cells is one of a smooth surface, whereas B-lymphocytes generally show multiple projections (Figure 1, 2).

T- and B-cell interaction is still relatively poorly understood, but it is clear that B-cells destined to make humoral antibody can interact and somehow be amplified in their response by so-called helper T-cells. Evidence has been presented for the existence of a second population of T-cells termed suppressor T-cells (26), which

Figure 1 Lymphocytes separated by Ficoll-Hypaque examined by scanning electron microscopy showing spectrum of lymphocyte surface morphology. T-cells more frequently have smoother surfaces with smaller numbers of microvilli than most of their more villous B-counterparts (Magnification X 14,600). (Photograph generously provided by Dr. Aaron Polliack, Sloan-Kettering Institute for Cancer Research, New York.)

may function as regulators or governors for the immune response. Macrophages may also play an important role in the immune response by processing and perhaps focusing various antigens (27, 28). In addition a T-cell-derived humoral factor (29–31) may contribute to or amplify various cellular reactions and it has been suggested that certain complement components may serve as activators of B-cells (32, 33).

The development of techniques that allow detection of T- and B-lymphocytes in blood and tissue fluids has stimulated investigation of changes in these two types of cells in numerous disease states in hopes of elucidating the origin of the malignant clone in the case of lymphoreticular malignancies or the mechanism of the defect in diseases thought to be of immunologic origin. The surface labels mentioned above have been very useful in this regard; however, before describing their alterations in disease a word of caution should be given regarding the many pitfalls which exist

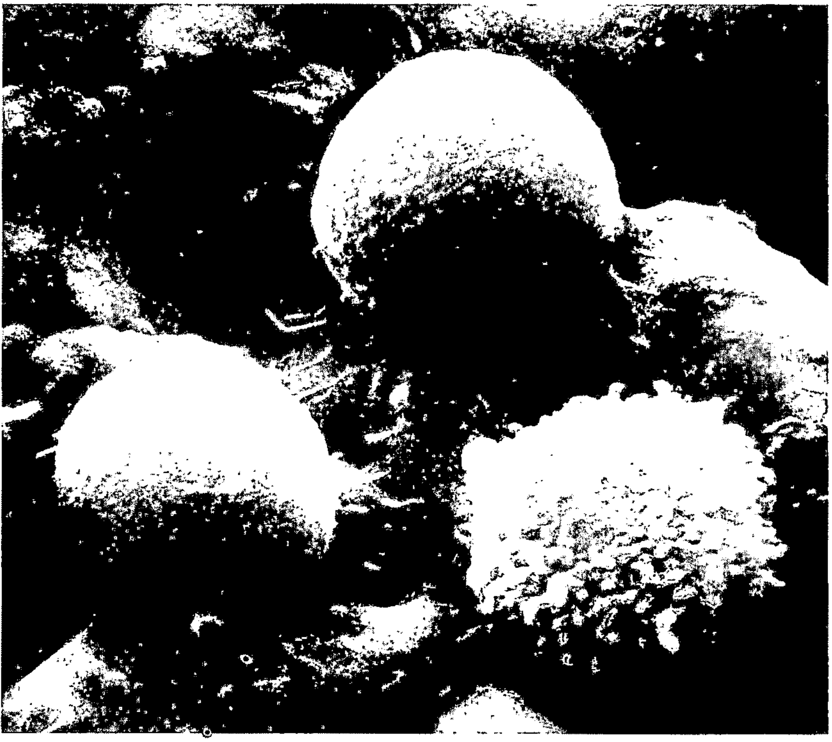


Figure 2 Scanning electron microscopic view of peripheral blood lymphocytes, showing both smooth and villous surfaces. The more smooth cells are generally felt to represent T-cells whereas B-cells show more prominent villous projections (Magnification X 13,200). (Photograph kindly provided by Dr. Aaron Polliack, Sloan-Kettering Institute for Cancer Research, New York.)

in this area of research. One problem relates to the presence of receptors for C3 (34) and the Fc portion of Ig (35) on the surface of monocytes and macrophages. Care must be taken not to include these cells in counts reported as B-lymphocytes. Several techniques used to avoid this error take advantage of the phagocytic properties of monocytes. They include labeling monocytes by allowing them to ingest polystyrene beads or yeast particles, and removing monocytes either by first inducing them to ingest iron filings and then extracting them in a magnetic field or by separating them over a surface to which they adhere such as glass or nylon wool. Unfortunately none of these techniques is completely specific. Another problem in distinguishing B-lymphocytes arises when antilymphocyte antibodies are present in the serum. When bound to the surface of T-lymphocytes they can cause these cells to be erroneously classified as surface immunoglobulin-bearing B-cells. Detection of autologous antilymphocyte antibody on T-cells by the fluorescent technique appears to be enhanced by handling the cells at lower temperatures (36, 37), by short incubation periods with fluorescein-labeled antisera, and by metabolic inhibition of the cells (36). It is also probable that antilymphocyte antibodies are heterogeneous and those of differing specificities or avidities behave differently in any given detection system. B-cells may also be overestimated when they are determined by the sum of cells bearing surface IgG, IgA, and IgM if the monospecific fluorescein-labeled antisera contain aggregates which adhere to the Fc or aggregate receptor of cells bearing surface Ig of another class. The possibility of the existence of cells that bear more than one type of surface Ig must also be considered when this technique is used. Other problems have arisen in situations in which the usual concurrence of receptors for C3, Fc, and surface Ig may break down (10, 38).

With respect to T-lymphocytes, small differences in technique have led to differing results from one laboratory to another. This is most apparent in the determination of T-cells by rosette formation with sheep erythrocytes. This test is dependent upon the duration of incubation of the red cells and lymphocytes. With short incubation times, normal values in the range of 30% are obtained while longer incubation results in normal values of approximately 70%. It is felt by many investigators that the early rosettes may represent a functional subclass of T-cells and thus vary independently from the late or total rosettes (39). This test is also affected by the presence of serum in the medium. Human type AB serum stabilizes the rosettes and raises the total number over that obtained if the test is done in its absence (40). Anti-T-cell sera also vary from one laboratory to another. Some sera are specific when used in the cytotoxic assay but not in the fluorescent method. In addition these heterologous antisera may differ in their specificity for different T-cell antigens and thus not be directly comparable.

In spite of these problems and the different techniques used by different investigators, a great deal of useful information has been accumulated on T- and B-lymphocytes in the past two years. Some authors have expressed T- or B-cells in terms of percentage of the lymphocytes, others in both percent and absolute numbers. It is apparent that these two methods of reporting data give insight into two different phenomena. The first bears on alterations in the balance of the cells active in cellular and humoral immunity; the latter bears on the magnitude of their potential compared to normal.

CHRONIC LYMPHOCYTIC LEUKEMIA

One of the first human disorders in which abnormalities of peripheral blood lymphocytes were extensively studied was chronic lymphocytic leukemia (CLL). The vast majority of cases appeared to show a proliferation of B-cells as documented by the presence of surface immunoglobulin (41, 42), predominantly of IgM type. Of great interest was the finding, confirmed by numerous investigators, that the surface Ig on CLL cells was of monoclonal type. A few examples of apparent absence of whole surface Ig or presence of light chains alone (43) have also been documented in association with CLL. In addition, the interesting finding was made by Preud'-Homme & Seligmann (44) that lymphocytes from certain patients with CCL contained surface immunoglobulin M, which was IgM rheumatoid factor capable of binding IgG to the cell surface receptors. This report emphasized the great care with which interpretation of data regarding apparently surface-bound Ig needs to be examined. Analysis of lymphosarcoma cells by Aisenberg & Bloch (45) showed considerably more cell-surface immunoglobulin than was generally seen with CLL cells.

Studies by Ross and co-workers (38) indicated that the presence of the C3 receptor on CLL cells varied independently of surface Ig receptors, confirming the microheterogeneity of surface characteristics among certain B-cell populations. We have recently studied a fascinating patient with CLL who appears to represent another exception to the general rule, in that peripheral blood lymphocytes showed simultaneous T- and B-cell markers (46). In this situation it was postulated that the malignant cell represented a hybrid of markers possibly due to derepression of genes which could add T-cell surface characteristics to a malignant stem B-cell. Sandilands et al (47) have also described a patient with a lymphoproliferative disorder in which cells had both T- and B-cell markers. They postulate, however, that the disorder was due to proliferation of a rare normal lymphocyte with both of these markers. Similarly, recent reports have documented relatively high proportions of CLL cells showing surface IgD as well as IgM (48). These findings amplify previous reports that in certain instances, particularly fetal lymphocytes, IgD and IgM may exist as surface markers on the same cells (49) and thus may be directly linked in ontogeny or a developmental sense. This is considered an important point since in general normal B-cells are felt only to express one class of surface immunoglobulin receptor (50).

MALIGNANT LYMPHOMA AND HODGKIN'S DISEASE

Malignant lymphomas and Hodgkin's disease represent a difficult area for analysis of T- and B-cell participation since, as one observer has recently emphasized, these forms of malignancy may involve a virtual civil war between T- and B-cells (51). Studies of tissue homogenates from malignant lymphomas have suggested a relatively high immunoglobulin content and therefore B-cell participation (52). In addition, enumeration of proportions of T- and B-cells in peripheral blood of patients with lymphoma appears to indicate an increase in B-cells (53, 54), particularly those bearing IgG or at least lymphocytes with detectable surface Ig.

Tissue B-cell hyperactivity in Hodgkins's disease is further suggested by recent reports of considerable in vitro IgG synthesis using cells obtained from spleens removed during surgery for staging of the disease (55). In some instances lymphocyte surface immunoglobulin detected in Hodgkin's disease may be adsorbed auto-antibody to lymphocyte surface antigens and not true intrinsic Ig antibody receptors. More studies of peripheral blood lymphocytes in such patients are necessary before the complex interactions of T- and B-cells can be unraveled in this disorder.

MULTIPLE MYELOMA

Malignant plasma cell myeloma has been of considerable interest as a disorder of the final common pathway of B-cell proliferation. Studies utilizing the mouse myeloma model have indicated that during induction of the tumor the proportions of normal peripheral blood lymphocytes bearing polyclonal Ig diminish, whereas proportions of lymphocytes showing monoclonal surface Ig of the same characteristics as the myeloma protein are increased (56, 57). Similar findings have been recorded in human multiple myeloma where decreases in proportions of peripheral blood lymphocytes bearing normal surface Ig have been recorded (58). Of great interest was the finding that patients with benign monoclonal gammopathy showed normal proportions of peripheral blood B-cells bearing normal polyclonal Ig. Somewhat different results have been recorded by Mellstedt et al (59) using the C3 receptor to enumerate B-cells. These authors reported an increase in both proportions and absolute numbers of lymphocytes binding to complement-coated erythrocytes. It is thus possible that in the case of plasma cell myeloma, surface Ig and C3 receptors on B-cells may vary in an opposite fashion.

Of considerable interest with respect to insight into the pathogenesis of multiple myeloma are the reports that the myeloma proteins may act as transplantation antigens in experimental animals. Preimmunization of normal mice with an isolated mouse myeloma protein appears to induce immunity to subsequent challenge with the plasma cell tumor capable of making that particular myeloma protein (60–62). These results suggest that considerable attention should now be directed at B-cell surface receptors as potential determinants of tumor immunity induction in man.

A high proportion of peripheral blood lymphocytes of patients with Waldenstrom's macroglobulinemia also have monoclonal surface IgM (63, 64). It has been postulated that in this disease an uninterrupted maturation of a malignant clone takes place, resulting in proliferation of the whole line from B-lymphocyte-to IgM-secreting plasma cells (63) rather than merely the beginning of the line as in CLL or the terminal cell in multiple myeloma.

ACUTE LYMPHOCYTIC LEUKEMIA

Studies currently available appear to indicate that acute lymphocytic leukemia in man is derived from a precursor with T-cell-like properties (53, 65). Studies by Shevach et al (66) using cultured cell lines derived from acute leukemic subjects showed absence of both C3 and surface Ig receptors. Of particular interest was the report of acute blast crisis occurring in a patient with chronic lymphocytic leukemia (67). In this particular instance monoclonal surface IgM with rheumatoid factor activity was detected on blast cells identical to that of the original CLL cells, indicating probable origin of both cell types from the same original clone.

IMMUNE DEFICIENCIES

Patients with X-linked agammaglobulinemia appear to show no detectable B-cells (68–74), whereas other individuals with agammaglobulinemia may show normal numbers of peripheral blood B-lymphocytes. In the latter instances specific defects of plasma-cell differentiation have been postulated (68, 70). Of interest have been the reports of normal B-cell values in subjects with IgA deficiency (71), again reflecting a block in the normal maturation and function of B-cells. Based on parallel B- and T-cell determinations in a group of patients with various immunodeficiencies, Dickler et al (13) have classified such patients into five categories. Surface Ig and the aggregate method were used to determine B-cells while T-cells were determined by the sheep cell rosette technique. One type showed normal proportions of B-cells but was felt incapable of differentiation to normal plasma cell components. A second type showed B-lymphocytes lacking easily detectable surface Ig, in which it was postulated that the B-cells were unable to synthesize or transport immunoglobulin. A third type of patient was noted to have large subpopulations of "null" cells, that is cells with neither T- nor B-cell markers, possibly representing precursor cells unable to differentiate into mature functionally active B-cells. A fourth group of patients showed large subpopulations of lymphocytes bearing both B- and T-cell markers, presumably multipotential precursor cells unable to complete final differentiation. The fifth group had total absence of all B-cells and represented individuals with sex-linked agammaglobulinemia. Studies such as these illustrate how application of current knowledge of cell surface marker characteristics may provide clear insight into the genesis of various immune deficiencies.

RHEUMATIC DISEASES

Systemic lupus erythematosus is a disease characterized by numerous immunologic abnormalities. Studies in this laboratory (72), recently confirmed by several other investigators (73–75), have clearly established that both the percentage and absolute number of T-cells are decreased and suggest that the decrease may be more marked in periods of increased disease activity. The percentage of cells bearing surface Ig has been shown to increase while the absolute number of these cells falls. The majority of patients also have antibody directed against their own lymphocytes (76, 77), which has been shown to react with both T- and B-cells (37, 78). The presence of these antibodies as well as the fact that the percentage of lymphocytes bearing the C3 receptor has been found reduced (72) has led to some caution in stating that there is actually a relative rise in B-lymphocytes in this illness. Study of synovial fluids in SLE has been scanty but it appears that the decrease in T-cells seen in the blood is mirrored in the synovial fluid (79). The decrease in T-cells coupled with the demonstration of defective cellular immunity and a hyperactive humoral immune system has sparked the idea that destruction of T-cells, particularly suppressor T-cells, may be one of the central pathogenic mechanisms involved in this disease.

Investigation of T- and B-cells in rheumatoid arthritis has provided less dramatic results. The percentage of T- and B-cells appears normal in most patients with rheumatoid arthritis, while the absolute numbers of both are decreased (75, 80). Some patients however appear to have a decreased percentage of T-cells. These patients tend to have more active disease with systemic involvement (22, 75). A longitudinal study done in our laboratory comparing clinical parameters with T- and B-cells has shown a significant correlation ($p < .01$) between the absolute number of T-lymphocytes and the number of swollen and painful joints (80). An inverse correlation was noted between the total number of lymphocytes bearing surface Ig (B-cells) and ring size. Other clinical parameters such as morning stiffness, walking time, and overall physician assessment showed no correlation with either percent or absolute numbers of T- or B-cells. Thus on the one hand a decrease in the percentage of T-cells appears to correlate with activity, while on the other hand increases in the absolute numbers of T-cells show the same type of correlation. Data on synovial fluid T- and B-cells are also confusing. B-cells have been reported to be increased (10), equal to (79, 81, 82), and decreased (83) as compared to peripheral blood, while T-cells have been found normal or increased (81–83). In one study a correlation was found between the presence of null cells and IgG-IgG complexes (82). Recently Van Boxel has examined the ratio of T- and B-lymphocytes in the synovial membrane of patients with rheumatoid arthritis and found it approximately equal to that in the blood (84). More data are needed before any conclusions can be drawn about the role of T- and B-cells in this disease.

Sjögren's syndrome is a pivotal disease between lymphoproliferative disorders and the rheumatic diseases. These patients appear to have an increased percentage of B-lymphocytes, particularly of those with surface IgM (85). T-lymphocytes are normal in the blood of most patients but are significantly decreased in others, especially those with co-existent rheumatoid arthritis. The high percentage of B-cells bearing IgM is intriguing since the most common lymphoid malignancy in these patients involves an IgM-secreting cell. As in rheumatoid arthritis, examination of tissues for T- and B-cells has revealed infiltration with cells of both types (85).

INFECTIONS

Data have begun to accumulate regarding alterations in T- and B-cells during the course of various viral infections. Infectious mononucleosis with its characteristic lymphocytosis and atypical lymphocytes has been studied by several groups. The report by Sheldon et al (86) suggests that atypical lymphocytes are derived from T-cells. Serial studies of typical patients during the course of the disease have revealed an initial elevation of the total numbers of both T- and B-cells (87). During the acute phases of the disease lymphocytotoxic antibody may appear (87–89). In some instances these antibodies appear to show specificity for subpopulations of B-lymphocytes (90).

A depression in proportions of T-lymphocytes has been reported during viral upper respiratory infections (39) and associated with other viral illness (91). Perhaps the most striking effects documented during viral infections have related to changes encountered in viral hepatitis. Marked depressions in both proportions and total numbers of T-cells have been recently recorded (92) during the early course of viral hepatitis. Of considerable interest in this latter study were the normal proportions and total numbers of T- and B-cells recorded among asymptomatic carriers of hepatitis-B antigen. Also of note were significant depressions of both proportions and total numbers of T-lymphocytes during the course of chronic active hepatitis (92). From these accumulated observations it seems clear that T-cell depression may be a feature commonly encountered during active viral infection.

Considerably more data are needed before a final general assessment of T- and B-cell alterations during various bacterial infections can be made. Initial studies (91) appeared to indicate some tendency for sequential depression of proportions of T-cells and elevations of B-cell percentages during the first week of various bacterial infections. However, notable exceptions or divergences are certainly present including the situation encountered during active lepromatous leprosy where a depression of cell-mediated immunity is often associated with marked hyperglobulinemia and auto-antibody formation. Many individual types of bacterial infections need to be examined before a general pattern becomes clear.

EFFECTS OF AGE

T- and B-cell markers as well as functional responses have been studied at the two extremes of life—in infancy and in old age. B-cell proportions in cord blood appear to be relatively normal (93). Of interest however are the elevated percentages of cells showing surface IgD in fetal samples (48). The percent and absolute numbers of T- and B-cells are also normal in pregnant women (94). Studies of T-cell function as estimated by mitogen response to phytohemagglutinin have been reported to be diminished in old age (95, 96). In addition, enumeration of T-cells by two groups of investigators has shown that proportions and absolute numbers of T-cells are reduced in subjects above 65 years old (97, 98). In both studies a considerable proportional increment in peripheral blood B-cell percentages occurred among older people. A decline in T-cells and T-cell function, particularly suppressor T-cells in conjunction with an increment in B-cells, might relate to the known increased incidence of auto-antibodies during old age (99, 100).

MISCELLANEOUS STATES

T- and B-cells have been examined among patients with Hashimoto's thyroiditis (101). Proportions of T-lymphocytes were reduced whereas no significant change was noted in B-cells. T-cells have been reported to be increased during Graves' disease (102).

Until recently, participation of lymphocytes in various body secretions has received little attention. Colostrum has now been shown to contain a small proportion of viable lymphocytes, at least 50% of which are T-cells (103). It is conceivable that such cells represent a source for lymphokines or T-cell products such as transfer factor, and are possibly capable of transfer to the nursing infant.

Studies related to inflammatory bowel disease have indicated a clear depression of both numbers and proportions of peripheral blood T-cells among patients with regional ileitis but not in ulcerative colitis (104). No clear relationship to previous course, extent of small bowel involvement, or concurrent therapy was documented. Liver involvement associated with ulcerative colitis was also associated with depression in percentages and total numbers of peripheral blood T-cells. Coupled with the lowered T-cells in chronic active hepatitis (92), these findings may suggest that several forms of liver damage in some way affect total numbers and proportions of peripheral blood T-cells.

The whole area of cancer immunology has afforded a fertile field to apply some of the concepts generated by dividing the functions of T- and B-cells. An early report by Stjernswärd et al (105) suggested that diminution in proportions of available T-cells after irradiation for mammary carcinoma might predispose the patient to subsequent metastases. Other reports (39) have also suggested that diminution in T-cell reserves may be associated with various malignancies and may prove useful as an indicator of poor host defense against the tumor. It seems likely that a clear understanding of T-cell kinetics may be important for understanding the immune response to various tumors.

Literature Cited

1. Raff, M. C., Sternberg, M., Taylor, R. B. 1970. *Nature (London)* 225:553–54
2. Pernis, B., Forni, L., Amante, L. 1970. *J. Exp. Med.* 132:1001–18
3. Raff, M. C. 1970. *Immunology* 19: 637–50
4. Klein, E. et al 1970. *Exp. Cell. Res.* 62:133–48
5. Rabellino, E., Colon, S., Grey, H. M., Unanue, E. 1971. *J. Exp. Med.* 133: 156–67
6. Unanue, E. R. et al 1971. *J. Exp. Med.* 133:1188–98
7. Papamichail, M., Brown, J. C., Holborow, E. J. 1971. *Lancet* 2:850–52
8. Bianco, D., Patrick, R., Nussenzweig, U. 1970. *J. Exp. Med.* 132:702–20
9. Basten, A., Miller, J. F. A. P., Sprent, J., Pye, J. 1972. *J. Exp. Med.* 135: 610–26
10. Mellbye, O. J., Messner, R. P., DeBord, J. R., Williams, R. C. Jr. 1972. *Arthritis Rheum.* 15:371–80
11. Dickler, H. B., Kunkel, H. G. 1972. *J. Exp. Med.* 136:191–96
12. Möller, G. 1974. *J. Exp. Med.* 139: 969–82
13. Dickler, H. B., Adkinson, N. F. Jr., Fisher, R. I., Terry, W. D. 1974. *J. Clin. Invest.* 53:834–40
14. Greaves, M. F. 1970. *Transplant. Rev.* 5:45–75
15. Bankhurst, A. D., Warner, N. L., Sprent, J. 1971. *J. Exp. Med.* 134: 1005–15
16. Crone, M., Kock, C., Simonsen, M. 1972. *Transplant. Rev.* 10:36–56
17. Grey, H. M., Colon, S., Solomon, A., McLaughlin, C. L. 1973. *J. Immunol.* 111:1923–25
18. Raff, M. C. 1969. *Nature (London)* 224:378–79
19. Greaves, M. F., Raff, M. C. 1971. *Nature (New Biol.)* 233:239–41
20. Cerottini, J. C., Nordin, A. A., Brunner, K. T. 1970. *Nature (London)* 228:1308–9
21. Gershon, R., Kindo, L. 1970. *Immunology* 18:723–37
22. Williams, R. C. Jr. et al 1973. *J. Clin. Invest.* 52:283–95

23. Fröland, S. 1972. *Scand. J. Immunol.* 1:269–80
24. Jondal, M., Holm, G., Wigzell, H. 1972. *J. Exp. Med.* 136:207–15
25. Wybran, J., Carr, M. C., Fudenberg, H. H. 1972. *J. Clin. Invest.* 51:2537–43
26. Allison, A. C., Denman, A. M., Barnes, R. D. 1971. *Lancet* 2:135–40
27. Unanue, E. R., Askonas, B. A. 1968. *Immunology* 15:287
28. Seeger, R. C., Oppenheim, J. J. 1972. *J. Immunol.* 109:255–61
29. Miller, H. C., Schmiege, S. K., Rule, A. 1973. *J. Immunol.* 111:1005–9
30. Bach, J. F. et al 1971. *Proc. Nat. Acad. Sci. USA* 68:2734–38
31. Watson, J. 1973. *J. Immunol.* 111: 1301–13
32. Dukor, P., Hartmann, K. U. 1973. *Cell. Immunol.* 7:349–56
33. Dukor, P. et al 1974. *J. Exp. Med.* 139:337–54
34. Huber, H., Polley, M. J., Linscott, W. D., Fudenberg, H. H., Müller Eberhard, H. J. 1968. *Science* 162:1281–83
35. LoBuglio, A. F., Cotran, R. S., Jandl, J. H. 1967. *Science* 158:1582–85
36. Kennedy, M., Messner, R. P. 1974. *Clin. Res.* 22:181A
37. Winfield, J. B., Winchester, R. J., Wernet, P., Fu, S., Kunkel, H. G. 1974. *Clin. Res.* 22:432A
38. Ross, G. D., Rabellino, E. M., Polley, M., Grey, H. M. 1973. *J. Clin. Invest.* 52:377–85
39. Wybran, J., Fudenberg, H. H. 1973. *J. Clin. Invest.* 52:1026–32
40. Bentwich, Z., Kunkel, H. G. 1973. *Transplant. Rev.* 16:29–50
41. Preud'Homme, J. L., Seligmann, M. 1972. *Blood* 40:777–94
42. Pernis, B. et al 1971. *Progr. Immunol. Int. Congr. Immunol.* 1:95–106
43. Grey, H. M., Rabellino, E., Pirofsky, B. 1971. *J. Clin. Invest.* 50:2368–75
44. Preud'Homme, J. L., Seligmann, M. 1972. *Proc. Nat. Acad. Sci. USA* 69: 2132–35
45. Aisenberg, A. C., Bloch, K. J. 1972. *N. Engl. J. Med.* 287:272–76
46. Hsu, C. C. S., Marti, G. E., Schrek, R., Williams, R. C. Jr. 1974. *Fed. Proc.* 33:731
47. Sandilands, G. P. et al 1974. *Lancet* 1:903–4
48. Fu, S. M., Winchester, R. J., Kunkel, H. G. 1974. *J. Exp. Med.* 139:451–56
49. Rowe, D. S., Hug, K., Forni, L., Pernis, B. 1973. *J. Exp. Med.* 138:965–72
50. Fröland, S. S., Natvig, J. B. 1972. *J. Exp. Med.* 136:409–14
51. DeVita, V. T. 1973. *N. Engl. J. Med.* 289:801–2
52. Stein, H., Lennert, K., Parwaresch, M. R. 1972. *Lancet* 2:855–57
53. Chin, A. H., Saiki, J. H., Trujillo, J. M., Williams, R. C. Jr. 1973. *Clin. Immunol. Immunopathol.* 1:499–510
54. Grifoni, V., Del Giaco, G. S., Manconi, P. E., Tognella, S. 1972. *Lancet* 1:848
55. Longmire, R. L. et al 1973. *N. Engl. J. Med.* 289:763–67
56. Yakulis, V., Bhoopalam, N., Schade, S., Heller, P. 1972. *Blood* 39:453–64
57. Bhoopalam, N., Yakulis, V., Costea, N., Heller, P. 1972. *Blood* 39:465–71
58. Lindström, F. D., Hardy, W. R., Eberle, B. J., Williams, R. C. Jr. 1973. *Ann. Intern. Med.* 78:837–44
59. Mellstedt, H., Jondal, M., Holm, G. 1973. *Clin. Exp. Immunol.* 15:321–30
60. Lynch, R. G. et al 1972. *Proc. Nat. Acad. Sci. USA* 69:1540–44
61. Hannestad, K., Kao M.-S., Eisen, H. N. 1972. *Proc. Nat. Acad. Sci. USA* 69:2295–99
62. Cosenza, H., Köhler, H. 1972. *Proc Nat. Acad. Sci. USA* 69:2701–5
63. Preud'Homme, J. L., Seligmann, M. 1972. *Blood* 20:777–94
64. Wernet, P., Feizi, T., Kunkel, H. G. 1972. *J. Exp. Med.* 136:650–55
65. Yata, J. 1970. *Clin. Exp. Immunol.* 7:781–92
66. Shevach, E. M., Herberman, R., Frank, M. M., Green, I. 1972. *J. Clin. Invest.* 51:1933–38
67. Brouet, J. C., Preud'Homme, J. L., Seligmann, M., Bernard, J. 1973. *Brit. Med. J.* 4:23–24
68. Cooper, M. D., Lawton, A. R., Bockman, D. E. 1971. *Lancet* 1:791–95
69. Siegal, F. P., Pernis, B., Kunkel, H. G. 1972. *Eur. J. Immunol.* 1:482
70. Gajl-Peczalska, K. J., Park, B. H., Biggar, W. D., Good, R. A. 1973. *J. Clin. Invest.* 52:919–28
71. Preud'Homme, J. L., Griscelli, C., Seligmann, M. 1973. *Clin. Immunol. Immunopathol.* 1:241–56
72. Messner, R. P., Lindström, F. D., Williams, R. C. Jr. 1973. *J. Clin. Invest.* 52:3046–56
73. Scheinberg, M. D., Cathcart, E. S. 1973. *Arthritis Rheum.* 16:566 (Abstr.)
74. Yu, D. T. et al 1974. *Arthritis Rheum.* 17:37–45
75. Utsinger, P. D., Bluestein, H. G. 1974. *J. Rheum.* 1:75 (Suppl.)
76. Mittal, K. K. et al 1970. *Nature (London)* 225:1255–56

77. Terasaki, P. I., Mattironi, V. D., Banaett, E. V. 1970. *N. Engl. J. Med.* 283: 724–28

78. Lies, R. B., Messner, R. P., Williams, R. C. Jr. 1973. *Arthritis Rheum.* 16: 369–75

79. Brenner, A. I., Scheinberg, M. S., Cathcart, E. S. 1974. *J. Rheum.* 1:75 (Suppl.)

80. Lindström, F. D., Messner, R. P., Williams, R. C. Jr. Unpublished data

81. Jacobs, L., Edgington, T., McGlamary, J., Johnson, J. 1974. *J. Rheum.* 1:75 (Suppl.)

82. Winchester, R. J., Siegal, F. P., Bentwich, Z., Kunkel, H. G. 1973. *Arthritis Rheum.* 16:138

83. Frφland, S. S., Natvig, J. B., Husby, G. 1973. *Scand. J. Immunol.* 2:67–72

84. Van Boxel, J. A. 1974. *J. Rheum.* 1:77 (Suppl.)

85. Talal, N. et al 1974. *J. Clin. Invest.* 53:180–89

86. Sheldon, P. J. et al 1973. *Lancet* 1:1153–55

87. Enberg, R. N., Eberle, B. J., Williams, R. C. Jr. 1974. *J. Infec. Dis.* 130:104–11

88. Mottironi, V. D., Terasaki, P. I. 1970. *Histocompatibility Testing.* Baltimore: Williams & Wilkins

89. Huang, S.-W. et al 1973. *J. Clin. Invest.* 52:1033–40

90. Thomas, D. B., Phillips, B. 1973. *Clin. Exp. Immunol.* 14:91–96

91. Niklasson, P. M., Williams, R. C. Jr. 1974. *Infec. Immunity* 9:1–7

92. DeHoratius, R. J., Strickland, R. G., Williams, R. C. Jr. 1974. *Clin. Immunol. Immunopathol.* 2:353–60

93. Lawton, A. R., Self, K. S., Royal, S. A., Cooper, M. D. 1972. *Clin. Immunol. Immunopathol.* 1:84–93

94. Brain, P., Marston, R. H., Gordon, J. 1972. *Brit. Med. J.* 4:448

95. Toh, B. H. et al 1973. *Clin. Exp. Immunol.* 14:193–202

96. Weksler, M. E., Hutteroth, T. H. 1974. *J. Clin. Invest.* 53:99–104

97. Diaz-Jouanen, E., Strickland, R. G., Williams, R. C. Jr. 1974. *Am. J. Med.* In press

98. Augener, W., Cohen, C., Reuter, S., Brittinger, G. 1974. *Lancet* 1:1164

99. Heimer, R., Levin, F. M., Rudd, E. 1963. *Am. J. Med.* 35:175–81

100. Hallgren, E. M., Buckley, C. C. III, Gilbertsen, V. A., Yunis, E. J. *J. Immunol.* 111:1101–7

101. Urbaniak, S. J., Penhale, W. J., Irvine, W. J. 1973. *Clin. Exp. Immunol.* 15:345–54

102. Aoki, N., Wakisaka, G., Nagata, I. 1973. *Lancet* 2:49

103. Diaz-Jouanen, E., Williams, R. C. Jr. 1974. *Clin. Immunol. Immunopathol.* In press

104. Strickland, R. G., Korsmeyer, S., Soltis, R. D., Wilson, I. D., Williams, R. C. Jr. 1974. *Gastroenterology* 67:570

105. Stjernswärd, J., Vánky, F., Jondal, M., Wigzell, H., Sealy, R. 1972. *Lancet* 1:1352

Copyright 1975. All rights reserved

SCANNING TECHNIQUES TO ASSESS THYROID NODULES

❖7142

James L. Quinn, III, M.D. and Robert E. Henkin, M.D.
Nuclear Medicine Section, Department of Radiology, Northwestern University—McGaw Medical Center, Chicago, Illinois 60611

INTRODUCTION

The first organ scanned using a radionuclide was the thyroid (1, 2). This procedure remains today an important diagnostic aid in the evaluation of patients with a palpable thyroid nodule. The thyroid scan serves as both a functional and structural evaluation of the gland. Indications for thyroid scanning generally include (3–5): 1. A mass in the neck or tongue which could be associated with the thyroid; 2. Evaluation of a substernal mass; 3. Evaluation of thyroid function after suppression and/or stimulation; 4. Search for occult thyroid malignancy; 5. Search for functioning thyroid metastases.

NORMAL ANATOMY

The normal thyroid scan usually has a butterfly appearance with the lobes tapering cephalad and becoming more rounded caudally. The radionuclide uptake is ordinarily homogeneous throughout the gland with thinning in the isthmus.

The pyramidal lobe, when seen, most frequently arises from the isthmus or adjacent portions of either thyroid lobe. It is more likely to be seen on the left and has been reported in as high as 35% of otherwise normal thyroid scans (6). When the thyroid is asymmetric the right lobe is usually larger. The thyroid scan is a poor determinant of absolute thyroid size because it is a two-dimensional representation. Ordinarily the lobes are approximately 4–5 cm in height and 2 cm in width.

IMAGING TECHNIQUES

Either a moving rectilinear scanner or a stationary camera device can be used to image the thyroid. With a pinhole collimator and employing radionuclides with gamma ray emissions in the 130–170 keV range the resolution of the camera is superior to any currently available scanning system.

An optimal *information density* for thyroid imaging is 2000–3000 counts per square cm. This is not practical unless either ^{99m}Tc pertechnetate or ^{123}I is used.

The thyroid scan is usually performed with the patient in a supine position and his neck extended. This decreases thyroid movement as well as helps to lift the thyroid away from interfering structures such as the sternum and the heads of the clavicles. Correlation of the scan findings with the physical examination of the neck is essential. All palpable masses must be carefully marked and outlined on either the camera or scanner thyroid images for proper interpretation.

NUCLIDES USED FOR THYROID IMAGING

Several nuclides have been used to image the thyroid. (Table 1). Today the most commonly used nuclide for thyroid scanning remains the first used, ^{131}I. ^{125}I enjoyed a brief era of popularity (7). It offered potentially better resolution and a reduced absorbed dose to the gland. ^{123}I, first proposed in 1966 (8), is not only the radioiodide but also the radionuclide of choice for thyroid imaging (9). ^{123}I is superior to ^{99m}Tc thyroid imaging: (*a*) when the uptake is low, (*b*) when one is searching for a substernal extension in the gland, and (*c*) when there are cold nodules present along the border of the gland. This cyclotron product, with a 13-hr physical half-life, is unfortunately not generally available except in large metropolitan areas.

^{99m}Tc pertechnetate has also been used for thyroid scanning (10–13). It became a popular scanning agent because of its low cost, its ready availability, the low radiation dose to the thyroid, and the high photon yield. It is handled differently than radioiodine by the thyroid in that it is trapped but does not undergo appreciable

Table 1 Nuclides used for thyroid imaging

Nuclide	Physical Half-Life	Imaging γ (keV)	β	Dose to Thyroid mrad/μCi	Comments
^{131}I	8 days	364	Yes	1500	Useful for uptakes and scanning. Best for whole body scanning for thyroid metastases.
^{125}I	60 days	35.5	No	1100	Useful for scanning only. Not good for substernal evaluation.
^{123}I	13 hr	159	No	15	Good for uptakes and scans. Iodide radionuclide of choice.
^{99m}Tc	6 hr	140	No	0.2	Useful for scanning. The TcO_4^--form images trapping only. Uptakes possible but difficult.
^{127}I	—	28.5	—	15	Fluorescent imaging of nonradioactive ^{127}I in gland using ^{241}Am, 60 keV source, and semiconductor detector.
^{75}Se	120 days	121 265 280	No	6	In the selenomethionine form has been used to identify malignancy in cold nodules. Uptake in benign autonomous thyroid tissue as well as in parathyroid compromises this technique. Nonroutine.
^{67}Ga	78 hr	93 185 296	No	0.27[a]	Concentrates only in poorly differentiated thyroid malignancies.

[a] Whole body. Thyroid dose not available.

organification and binding. It has been shown to be superior to ^{131}I thyroid imaging (14), but has technical drawbacks, especially in patients with low iodide uptakes, uptake in adjacent salivary tissue, vascular structures, and the esophagus (Figure 1). The normal uptake of pertechnetate in the thyroid gland at 30 min is 0.5–4.0% (13).

The abundant nonradioactive ^{127}I in the thyroid gland can be imaged using a technique developed at the University of Chicago (15). A beam, from a source of ^{241}Am, is directed at the thyroid gland. Its 60-keV photons knock a *K*-shell electron out of the ^{127}I atom. The resultant emitted cascade of characteristic radiations includes a 28.5-keV X ray which can be imaged using a lithium-drifted silicon semiconductor detector. This has special application in imaging the thyroid in children, pregnant females, patients whose iodide pool has been expanded by exogenous iodide, as well as in patients who are on thyroid replacement medication. This technique also offers a new parameter of thyroid physiology by supplying the total iodine content of the thyroid. Low iodine content has been noted in degenerative or hemorrhagic nodules, thyroid inflammation, adenomas, and carcinoma. Increased iodine content was found in Grave's and Plummer's thyrotoxicosis. Hashimoto's thyroiditis has shown an inconsistent pattern.

Several *tumor localizing* radiopharmaceuticals have been used in an attempt to increase the specificity of thyroid imaging. Among these are ^{75}Se selenomethionine

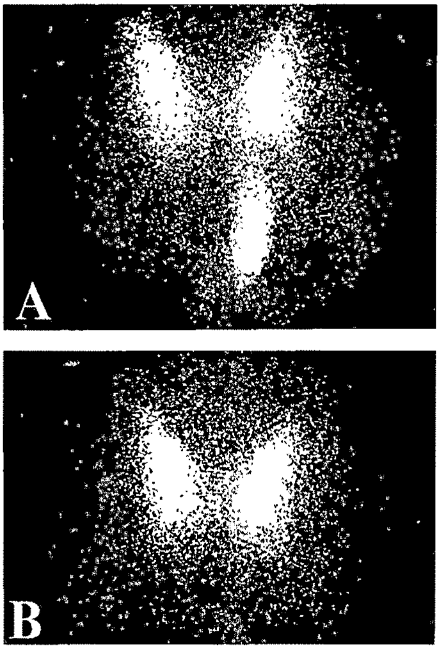


Figure 1 ^{99m}Tc pertechnetate images showing esophageal uptake (*A*), which disappears after swallowing a glass of water (*B*).

and ^{67}Ga citrate. Unfortunately, selenomethionine also localizes in areas of benign autonomous thyroid tissue as well as in the parathyroid gland and does not aid in the further differentiation of the thyroid nodule (16). ^{67}Ga citrate has been shown to concentrate in anaplastic or poorly differentiated tumors rather than in the more common papillary and follicular carcinomas (17).

The role of ultrasonography in differentiating benign cystic tumors, which are mixed or transonic, from solid tumors is as yet unsettled (18–20). Some thyroid carcinomas, as well as adenomas, show not only transonic qualities but also mixed echograms. Thermography has also been applied to thyroid nodule evaluation, carcinomas ordinarily being warmer than adenomas or cysts (18, 21).

THE PROBLEM OF NODULE DIFFERENTIATION

All this notwithstanding, we are still left with the problem of interpreting the pattern of radioiodide uptake in the thyroid of a patient with a single palpable nodule.

Clinically, carcinomas have been separated from adenomas because they (*a*) are firmer, (*b*) occur more frequently in younger people, and (*c*) don't decrease in size on thyroid hormone suppression. Though these are statistically valid, they are not absolutes upon which to make final patient decisions. This is where thyroid imaging can be of further assistance.

There are certain points which need to be remembered:

1. A nodule in a younger patient is more likely to be cancer than a similar nodule in an older patient, although the absolute incidence of thyroid carcinoma is higher in older patients.
2. A solitary thyroid nodule in a male is more likely to by thyroid carcinoma than a solitary nodule in a female, although the overall incidence of thyroid cancer is higher in females.
3. A nodule which grows rapidly is more likely to be carcinoma than a more slowly growing or nongrowing nodule, with the exclusion of hemorrhage into an adenoma.
4. Vocal cord paralysis in association with the palpable nodule has a higher probability of being a thyroid malignancy.
5. Cervical adenopathy, in association with a thyroid nodule, has a higher probability of carcinoma.
6. Nodules which disappear completely following suppressive hormone therapy are unlikely to be carcinoma.
7. Thyroid nodules which are "hot" on scans are unlikely to be malignant.
8. A single palpable nodule which on scan is shown to be one of many cold nodules in an adult is less likely to be malignant.

Indications for surgical removal of solitary nodules has been reviewed recently (22, 23).

It has been characteristic to refer to the uptake in a nodule in thermal terms, being "hot" (having greater activity and actually suppressing the surrounding thyroid), "cold" (having less than normal thyroid activity), or "warm" (having activity equal to the normal surrounding thyroid). This is probably more arbitrary than scientific but this is not the proper forum to undermine such established tradition.

Cold Nodules

The incidence of carcinoma in a solitary cold nodule has been reported to vary from 8–44% (24, 25); the generally agreed upon figure is approximately 20%.

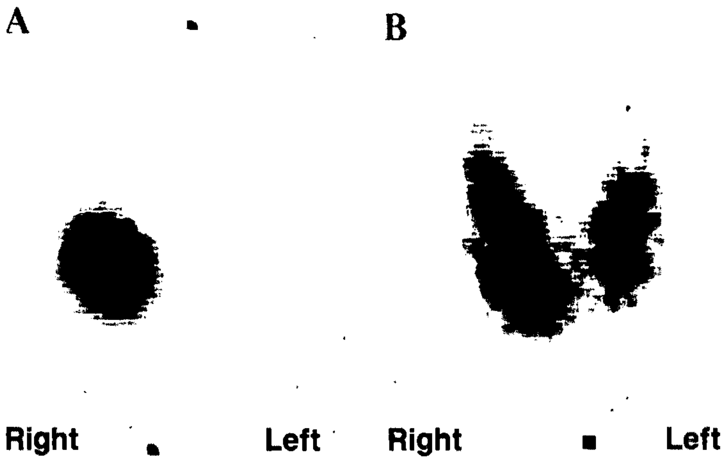


Figure 2 ^{125}I (*A*) scan in a 24 year old female showing the classic hot nodule suppressing the remainder of the gland, (*B*) ^{131}I scan: The gland is seen to function after TSH stimulation. Histologically the nodule was benign.

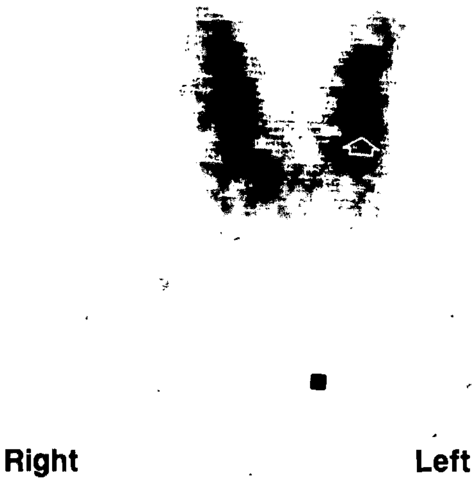


Figure 3 ^{123}I scan in a 25 year old male with a history of thymic irradiation showing a nondelineated or warm nodule in the left lobe which was a well differentiated papillary adenocarcinoma histologically.

Warm Nodules

The warm or more properly nondelineated nodule is a vexing problem. Oblique views are important in such cases and one may find that the palpable nodule which appears to function is indeed nonfunctioning and is located anterior or posterior to the normal functioning thyroid (26).

There is no autoradiographic evidence that carcinomas ever concentrate radioiodide as well as normal thyroid tissue. Therefore, the nondelineated or warm nodule most probably represents carcinoma interspersed with normal functioning thyroid tissue or a carcinoma of insufficient size to be detected by current imaging systems. Using ^{131}I the minimum detectable nodule had to be approximately 12 mm in size and have a decrease in radionuclide concentration of at least 20%. With camera magnification techniques employing pinhole collimation and using ^{123}I, a nodule as small as 4 mm can be identified on the thyroid images.

Hot Nodules

Hot nodules that are shown to suppress surrounding normal thyroid tissue are almost never malignant (Figure 2).

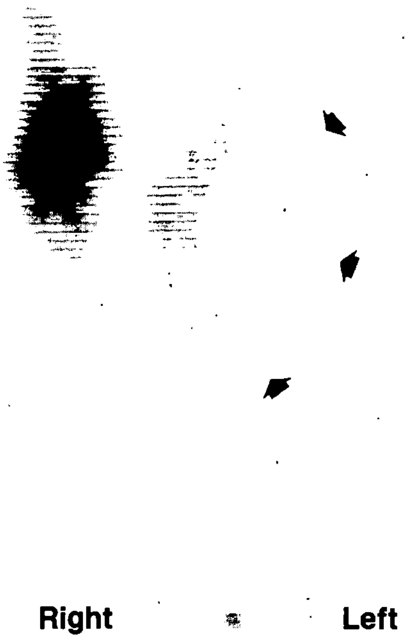


Figure 4 ^{125}I thyroid scan in a 53 year old female showing a cold nodule in the left lobe which histologically was a well differentiated follicular carcinoma.

Miller & Hamburger published excellent discussions regarding the interpretation of a hot or cold nodule on the thyroid scan (27, 28). Nonetheless it still remains an enigma. The exceptions to any rule of thumb are many.

It was suggested that nodules that are too small to be seen on the scan (Figure 3), or those which cause enlargement of a lobe without change in density, are unlikely to be malignant. Warm nodules fall into this category and probably have no higher incidence of malignancy in adults than found incidentally at autopsy, viz 4% (29). The major exceptions to this would be in patients with a history of previous irradiation to the neck. Sixty to eighty percent of patients with thyroid carcinomas have such a history (30).

Large nodules, which cause a slight replacement of normal tissue by virtue of their location along the periphery of the gland, are also seldom reported to be malignant (Figure 4).

[illegible]

carcinoma (Figure 5).

Medullary carcinomas, which are associated with pheochromocytomas and produce excessive thyrocalcitonin, present as cold nodules on thyroid scans usually at the junction of the upper and middle third of the thyroid gland (31).

DOUBLE TRACER SCANNING

In a patient with suspected thyroid carcinoma, we prefer to do an initial scan using ^{99m}Tc pertechnetate and then follow this, if necessary, with an ^{123}I scan. The reason for this is that we as well as others have demonstrated the presence of thyroid malignancies that are able to trap but not organify iodide. The pertechnetate scan demonstrates these to be functioning nodules. More importantly, extrathyroidal nodal metastases may be demonstrated which are not seen with radioiodide scanning (Figure 6). However, the finding of a thyroid nodule that traps but does not organify is nonspecific and was found in only 3% of dual radionuclide studies (9, 32–34).

From a diagnostic standpoint, the potential for demonstrating nodal metastases with pertechnetate is useful. However, the radioiodine scan remains essential if there is any consideration of radioiodine therapy.

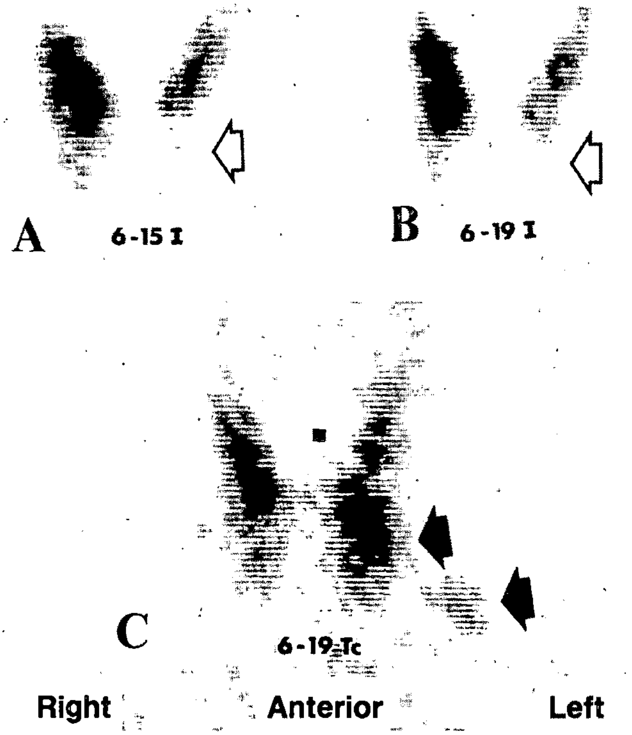


Figure 6 ^{125}I (*A*), ^{131}I (*B*), and ^{99m}Tc pertechnetate (*C*) thyroid scans in a 23 year old female with a palpable nodule in the lower pole of the left lobe of thyroid which was cold with radioiodide and warm with pertechnetate. This papillary adenocarcinoma had the ability to trap but not organify iodine. A metastasis to a cervical node is also identified only on the pertechnetate scan.

SUMMARY

Despite its nonspecificity, the thyroid scan remains an important diagnostic aid in the evaluation of the patient with a solitary thyroid nodule. Dual radionuclide thyroid imaging using ^{99m}Tc pertechnetate and ^{123}I offers more useful information than either alone. The role of complementary diagnostic modalities using ultrasonography and thermography looks promising but is as yet unclear.

Literature Cited

1. Cassen, B., Curtis, L., Reed, C., Libby, R. 1951. *Nucleonics* 9:46–50
2. Allen, H. C. Jr., Libby, R. L., Cassen, B. 1951. *J. Clin. Endocrinol.* 11:492–511
3. Kuhl, D. E. 1963. *Radiol. Clin. N. Am.* 1:101–14
4. Spencer, R. P. 1967. *N.Y. State J. Med.* 67:1139–43
5. Quinn, J. L. III, Behinfar, M. Radioisotope scanning of the thyroid. 1967. *J. Am. Med. Assoc.* 199:920–24
6. Izenstark, J. L., Forsaith, A. L., Horwitz, N. H. 1969. *J. Nucl. Med.* 10:519–24
7. Endlich, H., Harper, P., Beck, R., Siemans, W., Lathrop, K. 1962. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 87:148–55
8. Myers, W. 1966. *J. Nucl. Med.* 7:390
9. Atkins, H. L., Klopper, R. M., Lambrecht, R. M., Wolf, A. P. 1973. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 117:195–201
10. Andros, G., Harper, P. V., Lathrop, K. A., McCardle, R. J. 1965. *J. Clin. Endocrinol.* 25:1067–76
11. Sanders, T. P., Kuhl, D. E. 1968. *Radiology* 91:23–26
12. Atkins, H. L., Richards, P. 1968. *J. Nucl. Med.* 9:7–15
13. Atkins, H. L. 1971 *Semin. Nucl. Med.* 1:345–55
14. Strauss, H. W., Hurley, P. J., Wagner, H. N. Jr. 1970. *Radiology* 97:307–10
15. Hoffer, P. B., Gottschalk, A., Refetoff, S. 1972. *Curr. Probl. Radiol.* 2:#1
16. Thomas, C. G. Jr., Pepper, F. D., Owen, J. 1969. *Ann. Surg.* 170:396–408
17. Kaplan, W. D. et al 1974. *J. Nucl. Med.* 15:424–27
18. Planiol, T., DeRycke, G., Floyrac, C. 1972. *Corse Mediter. Med.* 18:81–87
19. Blum, M., Goldman, A. B., Herskovic, A., Hernberg, J. 1972. *N. Engl. J. Med.* 287:1164–69
20. Miskin, M., Rosen, I. B., Walfish, P. G. 1973. *Ann. Intern. Med.* 79:505–10
21. Samuels, B. I. 1972. *Radiology* 102:59–62
22. Green, W., Senturia, H., Packman, R., Richards, F. 1972. *J. Am. Med. Assoc.* 221:1265–68
23. Brooks, J. R. 1973. *Am. J. Surg.* 125:477–81
24. Johnson, P. C., Beierwaltes, W. H. 1955. *J. Clin. Endocrinol.* 15:865
25. Shimaoka, K., Sokal, J. E. 1964. *Arch. Intern. Med.* 114:36–39
26. Charkes, N. D. 1971. *Semin. Nucl. Med.* 1:316–33
27. Miller, J. M., Hamburger, J. I. 1965. *Radiology* 84:66–74
28. Miller, J. M., Hamburger, J. I. 1965. *Radiology* 85:702–10
29. Mortensen, J. D., Woolner, L. B., Bennett, W. A. 1955. *J. Clin. Endocrinol.* 15:1270–80
30. Winship, T., Rosvoll, R. V. 1961. *Cancer* 14:734–43
31. Wolf, J. H. et al 1973. *N. Engl. J. Med.* 289:437–41
32. Steinberg, M., Cavalieri, R. R., Choy, S. H. 1970. *J. Clin. Endocrinol. Metab.* 31:81–84
33. Usher, M. S., Arzoumanian, A. Y. 1971. *J. Nucl. Med.* 12:136–37
34. Shambaugh, G. E. III, Quinn, J. L. III, Oyasu, R., Freinkel, N. 1974. *J. Am. Med. Assoc.* 228:866–69

Copyright 1975. All rights reserved

OXYGEN TOXICITY

❖7143

Walter G. Wolfe, M.D. and William C. DeVries, M.D.
Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710

From the greater strength at vivacity of the flame of a candle, in this pure air, it may be conjectured, that it might be peculiarly salutary to the lungs in certain morbid cases when the common air would not be sufficient to carry off the phlogistic putrid effluvium fast enough. But perhaps we may also infer from these experiments that though pure dephlogisticated air might be very useful as a medicine it might not be proper for us in the usual healthy state of body: for as a candle burns out much faster in dephlogisticated than in common air so we might as may be said live out too fast and the animal powers be too soon exhausted in this pure kind of air. A moralist at least may say that the air which nature has provided for us is as good as we deserve.

Priestly 1775 (1)

Experimentation with the biological role of oxygen began immediately following its isolation in 1772. Lavoisier, who was the first man to describe the true nature of the interchange of gases in the lung, referred to the "incendiary action" upon the lungs of animals exposed to this gas. In the nineteenth century, Paul Bert recorded the first formal account of oxygen toxicity restricted to hyperbaric pressures (2). J. Lorrain Smith is credited with first describing the pathology of pulmonary oxygen toxicity (3, 4). As early as 1897, he recognized that the inherent toxicity of oxygen might limit its clinical application. The importance of this toxic effect of oxygen on the lung has not been appreciated by physicians engaged in patient care until the last 15–20 years. The toxic effects of high partial pressure of inspired oxygen upon the eyes of newborn infants were the first clinical evidence of the dangerous effect of oxygen administered at atmospheric pressure (5, 6).

Although oxygen had been used clinically for the treatment of hypoxemia for the last 30–40 years, pulmonary oxygen toxicity became a serious clinical reality with the development of efficient mechanical ventilators capable of delivering a high concentration of oxygen.

Excellent reviews of toxic effect of oxygen have been published by Clark & Lambertsen in 1971 (7) and Winter & Smith in 1972 (8). A review of the cellular mechanisms in oxygen toxicity was published in 1968 by Haugaard (9), and Bean in 1945 published a classical review of the effect of oxygen at hyperbaric pressures (10). This review attempts to highlight and focus on topics that may be relevant to the clinician.

MORPHOLOGY AND PHYSIOLOGY

The most important alterations known to occur in the respiratory tract of man and animals are changes in the airway structures, vascular bed, and lung tissue. One may also find exudates and bloody fluid appearing in the pleural cavity. Changes in airway structures include atelectasis, the accumulation of leukocytes, erythrocytes, macrophages in alveoli, and formations of membranes on the alveolar wall. Changes in the vascular bed include hyperemia, conglomeration of blood cells in capillaries, and capillary proliferation. Changes in lung tissue include peribronchial and perivascular edema and accumulation of fluid in cells and interstitium. Electron microscopy reveals swelling of mitochondria in the alveolar epithelium. These pathologic changes are observed for the most part in the terminal phase of oxygen toxicity.

At present, it appears clear that two types of pulmonary changes arise in response to oxygen, those associated with low dosages and those associated with high dosages. Low dose exposures are accompanied by proliferative changes in the endothelium and epithelium, and produce what is considered to be a chronic, often nonlethal, type of oxygen toxicity. In contrast are the exudative changes seen after exposure to high concentrations which result in acute oxygen toxicity. Acute changes are characterized by Smith (4) as having the histologic features of atelectasis, pulmonary capillary congestion, interstitial and alveolar edema, capillary degeneration, and hemorrhage. Studies by Kistler (11) using electron microscopy to evaluate the above changes in rats revealed that once gross evidence of toxicity occurred in the lung, there were signs of capillary endothelial, but no alveolar endothelial, damage. The earliest change was a twofold increase in the width of the interstitial space as a result of edema with only a small quantity of cellular infiltration. After 72 hr, this progressed to extensive destruction of capillaries, widening of the interstitial space, and marked cellular infiltration. It was not until this late stage that the alveolar changes of cellular swelling were first observed. In studies performed on the primate lung, the *endothelium* also showed changes earlier than the *epithelium* (12, 13).

In rats breathing oxygen, no observable changes in lung structure were found at 6 to 24 hr. However, after 48 hr exposure, the air-blood barriers became thickened by accumulation of interstitial fluid with early destructive changes of the capillary endothelium. After 72 hr there was marked destruction of the pulmonary capillaries; the interstitial space had doubled in thickness and contained leukocytes, thrombocytes, and other cells. At this point, 65% of the alveoli were filled with an exudate containing leukocytes, erythrocytes, and macrophages with fibrin strands (11).

Exposure of monkeys to 90 to 100% oxygen for up to 12 days resulted in 50% mortality from oxygen toxicity within seven days while the remainder became adapted and survived beyond that time. Cyanosis occurred immediately following exposure of this group to air and was thought to be secondary to the alveolar proliferative lesions. Within two days after exposure to the oxygen-enriched environment there was a slight swelling of endothelial cells and a small amount of interstitial fluid, although the morphometric measurements were within normal range. After four days there was considerable destruction of alveolar cells, largely of the type I membranous pneumatocytes. There was an increased thickness in the air-blood barrier, largely due to an increase in the interstitial fluid volume. After seven days, hyperplasia of the type II alveolar epithelial cells, granular pneumatocytes was evident, and there was a marked increase in endothelial thickness. Type I cells were almost completely destroyed by this time. The endothelium differed from region to region with gross variations in thickness. There was evidence of swollen interstitial spaces which were altered by increasing cellularity composed of fibroblasts and leukocytes (12, 13). Studies by Kistler (11) and Kapanci (13) demonstrated the differences between response to oxygen in the rat and monkey. Although endothelial changes were produced in both, the rat had a greater loss of capillary surface area associated with shorter survival time and showed little of the proliferative phase of acute oxygen toxicity.

Hyaline membranes seen in oxygen toxicity probably represent a combination of serous and cellular exudation following the initial critical endothelial damage. Hyaline membranes are therefore more often seen in oxygen toxicity when survival time is prolonged than when survival time is short.

In monkeys, the acute exudative phase lasts approximately 12 days. Rats manifested an exudative change at 40–70 hr. Whether an animal develops exudate or proliferative response depends not only on the P_{IO_2} and the species, but also on strain, age, and individual susceptibility. Again, it is emphasized that many of the changes of oxygen toxicity are nonspecific. Thus, septal edema, hyaline membranes, and hyperplasia of the alveolar lining cells may be produced by radiation, nitrogen oxide, pneumonia, or cytotoxic drugs. In a clinical situation where respiratory failure follows lung injury, separation of pneumonia, injury, and effect of oxygen poisoning is difficult to determine. Pratt has reported capillary proliferation in human lungs following oxygen therapy but this is not evident in animal studies (14).

The effect on the lung of 95% oxygen at normal barometric pressure in the awake, spontaneously breathing dog and in anesthetized dogs mechanically ventilated demonstrated the lung-body weight ratio to be highest in the spontaneously breathing dogs. Lung weights in both groups were greater than for the control animal. Histologic studies demonstrated marked perivascular edema in both groups. However, the effect of breathing 95% O_2 appeared to be less in the mechanically ventilated dogs, as demonstrated by the lower lung weights and smaller decrease in Pa_{O_2}. Lung biopsies were done in one group and the animals allowed to recover. Significant recovery from oxygen poisoning as measured by the disappearance of perivascular edema was not evident until approximately ten days following resumption of breathing room air. As evident on microscopy, the severity of the edema at 60–64 hr is less than that seen at 72 hr. Recovery was essentially complete at ten days, although in some sections traces of perivascular edema could be identified. At ten days prior to sacrifice in both mechanically ventilated and the spontaneously breathing dogs, the Pa_{O_2} had returned to normal (15).

The direct effect of oxygen on the lung was studied by Penrod (16) who ventilated cats with a double-lumen catheter at atmospheric pressure. One lung was exposed to oxygen and the other to room air. After several hours the oxygen ventilated lung showed severe changes but the air ventilated lung was normal. Studies using a tracheal divider to evaluate the effects of oxygen poisoning in dogs at atmospheric pressure on the mucociliary transport demonstrated a marked decrease in clearance in the oxygen breathing lung. Histologic changes were marked in the oxygen-exposed lung while changes noted in the lung exposed to air were minimal. Therefore, it appears that the partial pressure of the inspired oxygen is probably the most critical factor in producing pulmonary oxygen toxicity (17).

Animal experiments to investigate morphologic changes in lungs exposed to high partial pressures of oxygen in the shunted and unshunted dogs have been of interest. Miller (18) created a large transpulmonary shunt resulting in arterial oxygen tensions of 29 to 58 mm Hg in dogs breathing 100% O_2. These cyanotic dogs and control animals were exposed to 98–100% oxygen for 48 hr and then sacrificed. Although no detailed description of pathologic changes or methods of quantification were given, early evidence suggested that an increase in the alveolar-arterial O_2 difference (A-aD O_2) did not provide protection at normobaric pressures of oxygen. A confirmatory study was done by Ashbaugh (19) when he diverted inferior vena cava blood into the left atrium in eleven animals. These animals were exposed to 540–580 mm Hg partial pressure of oxygen until death. Daily measurements of blood gases, arterial oxygen saturation, systemic and pulmonary arterial pressures, cardiac output, and serum enzymes were made. There were no significant differences between survival times of both groups of dogs, the mean being approximately eight days for both. The study confirms the work of Miller (18) and indicates again that a large alveolar-arterial O_2 difference does not alter development of oxygen toxicity at atmospheric exposures.

Thomas (20) performed a complicated set of experiments in which the left pulmonary artery was ligated and the bronchial arteries were allowed to form active collateral circulation in the left lung. He then compared this experiment with one in which the right middle lobe pulmonary artery was ligated and an anatomic right-to-left shunt was formed. He found that during oxygen exposure the severity of the pulmonary changes at a constant Pa_{O_2} was related primarily to the amount of hyperoxemia in the blood perfusing the lungs. The morphological fact that the capillary endothelium is damaged much earlier than the alveolar epithelium gives credence to the possibility that oxygen tension in the perfusing blood may have some part in the pathophysiology of pulmonary oxygen toxicity.

Brauer (21) studied the effect of acclimatization to altitude in rats subsequently exposed to O_2. There was a threefold increase in survival in the altitude group compared to the controls. Despite controversy surrounding the pulmonary capillary versus the alveolar oxygen concentration, it is generally thought that the features of oxygen toxicity in the lung are dependent mainly upon the P_{IO_2} which is modified by individual and species variations. However, this argument is academic due to the impossibility of preparing the model necessary for such a study. The alveolo-capillary membrane is so permeable to oxygen that an adequate oxygen gradient could not be maintained long enough to prove either hypothesis.

BIOCHEMISTRY

Oxygen poisoning, regardless of organ system involved, is fundamentally a biochemical phenomenon of great complexity. As mentioned earlier, a complete summary of the literature was done by Haugaard (9) who prefaced his discussion by saying

> although oxygen is necessary for production of energy and survival of all aerobic cells, it is also a universal cellular poison. It is only because cells in the course of evolution have developed a special defense mechanism that life as we know it has been able to flourish freely. In a sense, the study of oxygen toxicity is a study of the ways in which the organisms protect themselves against the oxidizing potential of molecular oxygen.

Attempts have been made to protect against oxygen toxicity by modifying specific biochemical events. These have not yet proven successful, but if a truly protective mechanism could be found, it would be a significant step forward and would obviously have to come from a very basic investigation in this field. Most of the research of the biochemistry of oxygen toxicity has dealt with isolated enzyme systems and tissue preparations in vitro. Extrapolation to the intact organism must be made with caution.

Interestingly, the tissue with the highest intracellular oxygen tension is the lung where the alveolar and capillary cell Po_2 are a function of the partial pressure of alveolar gas, the O_2 tension in the blood, and the metabolic rate in the cells. It is not difficult to understand that under normobaric conditions the lung is the first organ to manifest oxygen damage (22).

In basic research concerning oxygen poisoning, oxygen has been found to be a potent enzyme inhibitor. The most attention has been given to enzymes containing sulphydryl groups. Several mechanisms for the oxidation of sulfhydryl groups, such as lipoic acid or coenzyme A, have been proposed but it is not yet known which pathways are involved. One view concerns the formation of free radicals and is supported by the observation that trace metals are essential for SH group oxidation. Another view supports the involvement of dual roles played by oxidation of glutathione, which allows it to react with the sulfhydryl groups of the enzymes, inactivating them, resulting in the formation of H_2O_2 which must then be degraded by catalyst. Finally, it has been demonstrated that reduced glutathione may serve to reactivate some oxidized sulfhydryl enzymes.

Inactivation of SH enzymes can have a profound effect as several such enzymes are involved in the tricarboxylic acid cycle and glyceraldehyde phosphate dehydrogenase is susceptible to oxidation. In cellular respiration, flavoproteins are vulnerable and Chance (23) demonstrated that oxygen interferes with the progression of electron transport in this system. Oxygen may interfere with oxidative phosphorylation of the mitochondria and may cause altered metabolism of glutamate and gamma amino butyric acid (GABA) which may be fundamental to the effect of oxygen at high pressure upon the central nervous system. Activation of the now clinically significant prostaglandin system may be important in explaining some of the biochemical pathways involved in pulmonary oxygen toxicity.

Metabolic and biochemical studies reveal decreased oxygen uptake by cell fractions, a reduction in the intracellular level of high energy phosphates, apparent inactivation of some enzymes containing sulfhydryl groups, and release into cytoplasm of enzymes normally confined to lysosomes (24, 25). Investigations have failed to indicate, however, whether any of these abnormalities is a cause rather than a result of the primary intracellular events that culminate in oxygen toxicity. Depending upon the experimental model, a variety of pharmacologic protectants have been identified. These, in general, delay but do not prevent the occurrence of oxygen toxicity. Among the protectants elucidated in various experimental models are vitamin E, succinate, GABA, glutathione, disulfiram, and certain monamine oxidase inhibitors. None of these agents has yet come into general clinical use in order to permit safe employment of oxygen in dosages that would otherwise be toxic.

Many oxidative reactions, both enzymatic and nonenzymatic, are known to cause the bivalent reduction of oxygen to hydrogen peroxide. Hydrogen peroxide is a reactive species and its accumulation inside cells would certainly lead to the death of these cells. Hydrogen peroxide was thus considered to be the agent of oxygen toxicity and catalases and peroxidases were thought to constitute the defenses which permitted aerobic life by preventing the accumulation of this reactive compound. This theory, though meritorious, is not entirely satisfactory because aerobes have been described which lack catalase and anaerobes have been described which contain this enzyme. In addition, the hydrogen peroxide theory of oxygen toxicity is incomplete in that it fails to consider reduction products of oxygen, which are even more reactive and more inimical to life than is hydrogen peroxide.

Oxygen exhibits a distinct preference for reactions in which one electron is transferred at a time. Reduction of oxygen often proceeds in univalent steps so that the first reduction product of oxygen, at neutral pH, will be O_2^-. This is called the superoxide radical and because of its reactivity and fleeting lifetime, it was for too long a period considered to be of concern only to radiation chemists, who generated it with bursts of ionizing radiation and who explored its properties during the milliseconds of its existence.

Recently, the superoxide radical has been of interest to biologists because this reactive radical is generated by a wide variety of biochemical events (26, 27). There are oxidative enzymes which, in their normal functioning, generate substantial amounts of superoxide radical and there are numerous nonenzymatic oxidations of materials, found inside cells, which also produce superoxide radical. Therefore, it follows that oxygen metabolizing cells will be exposed to a flux of the superoxide radical whose reactivity, if unopposed, would destroy these cells.

The primary defense appears to be an enzyme which catalyzes the reaction: $O^-_2 + O^-_2 \longrightarrow 2H + H_2O_2 + O_2$, and which has been named superoxide dismutase (28). This enzyme is ubiquitous among oxygen-metabolizing cells but is lacking in obligate anaerobes. Superoxide dismutase is an enormously efficient catalyst which operates at rates approaching the theoretical limit set by rates of diffusion. In microorganisms, it has been shown that raising the concentration of oxygen induces increased accumulation of this enzyme and further, that cells which contain high levels of superoxide dismutase are more resistant to the lethality of hyperbaric oxygen than comparable cells which contain lower levels of this enzyme. Mammalian cells are probably just as dependent upon superoxide dismutase for their ability to tolerate exposure to oxygen as are the microorganisms thus far investigated.

ATELECTASIS, ABSORPTIVE COLLAPSE, SURFACTANT, AND MUCOCILIARY TRANSPORT

In animals and man exposed to 100% oxygen, there is a phenomenon called absorption collapse which occurs during the breathing of oxygen. Evidence that this event is an initiating factor in the development of acute pulmonary oxygen toxicity is tenuous. Its effect on surfactant production and its contribution to the development of pneumonia may be important.

Collapse of alveoli during breathing of oxygen has been demonstrated. This collapse occurs at reduced lung volumes, and at normal lung volumes in the absence of coughing and sighing. It is evident, therefore, that in patients with severe pulmonary disease which includes narrowing of the airways resulting from mucous retention, inflammation, and other conditions, collapse may occur more readily when breathing oxygen than air.

Almost all the literature which describes pulmonary damage caused by oxygen also describes gross atelectasis. It has been found that the degree of consolidation increases with the amount of delay between terminal apnea and fixation of the lung. Atelectasis may therefore result from postmortem rather than antemortem changes. Consistent with this hypothesis has been a demonstration that in terminal stages of oxygen toxicity, apnea may precede cardiovascular collapse (29). It may be observed that animals breathing 95% oxygen usually underventilate. One may conclude that collapse of the lung which is present in the later stages of pulmonary oxygen toxicity is a result, not an initiating effect, of oxygen toxicity (15).

When 100% oxygen is breathed and gas is trapped within the airway spaces due to distal airway closure, it is absorbed rapidly into the blood which in turn results in atelectasis. The lack of nitrogen, a slowly absorbed gas, rather than the presence of oxygen is the key factor. Edmonds (30) demonstrated in dogs with ligation of one pulmonary artery that if the lung was inflated with nitrogen and atelectasis was prevented, surface tension extracts from lung remained normal.

The importance of surface tension in the lung was first demonstrated by Van Neegaard in 1929. Studies on the effect of oxygen in hyperbaric or atmospheric pressure on lung surfactant are controversial, some showing reduced activity and others showing normal activity (31, 32). Many of the chemical assays are open to serious question, particularly those involving whole lung extracts. The interpretation of in vitro measured changes of surfactant is difficult since many factors, such as contamination by serum, fibrinogen, and perhaps mechanical factors, are known to have an effect on the results.

Pulmonary edema is thought to result from mechanical changes in the lung caused by increased surface tension in the edematous parts. Although it is evident that there are measurable changes in surface tension in the late stage of acute pulmonary oxygen toxicity, it is still not clear whether these are the results of direct effects or indirect effects of the consequences of the changes in the lung induced by oxygen. In anesthetized dogs ventilated through a double-lumen catheter at 1 atm pressure, one lung with air and the other with 100% oxygen, there was a greater diminution of surfactant activity in the oxygen-ventilated lung than in the room-air-ventilated lung (33). A similar correlation existed with regard to reduction in the number of type II alveolar cells (believed to be the lung source of surfactant) and the number of lamella contained in these cells. This work should be contrasted with the results reported by Kapanci (13), which demonstrated little reduction in the number of type II cells in the conscious monkey. Although hyperplasia of type II cells in the later stage of acute pulmonary oxygen toxicity is documented, any relation between such changes and reduced surface tension is speculative. Similar studies in which one lung was ventilated with 100% oxygen and the other with air demonstrated no difference in the ^{32}P uptake in either lung, which was indicative of the synthesis of phospholipid in the lung. Surface tension measurements were also normal in both lungs (17). Similar findings were reported by Coalson (34) with regard to surface tension, but morphological changes in the air-ventilated lung suggested again a blood borne factor in pulmonary oxygen toxicity.

The role of the respirator in oxygen toxicity is controversial (35, 36). Most studies reveal no changes between the ventilated and nonventilated animal lungs and those exposed to oxygen. Nash (35), in a study using angora goats, found no differences in those treated with or without mechanical ventilators. Evidence at this time indicates that mechanical positive pressure ventilation, with physiologic inspiratory pressures, does not by itself cause morphologic pulmonary alteration, and those changes seen can be attributed to the high concentration of oxygen delivered.

Laurenzi (37) first called attention to the effect of oxygen on tracheomucous flow. He demonstrated that any deviation from ambient oxygen tension in the inspired air had an adverse effect on mucous flow. In dogs ventilated with air from 24 to 36 hr there was a two- to threefold decrease in mucociliary clearance whereas those ventilated with 100% oxygen had as much as a fivefold delay in clearance. Similar findings were observed in the same study in dogs ventilated through a divided tracheal catheter with one lung receiving 100% oxygen and the other air. Ventilation with high concentration of oxygen significantly decreased mucociliary transport. Previous studies and observations on ciliary cells grown in tissue culture seemed to indicate that high oxygen tension affects the mucociliary blanket or the mucous glands rather than the cilia directly. Available alveolar surfactant and the ability of the lung to produce surfactant were normal in the studies (17).

PULMONARY CIRCULATION

Studies in rats have demonstrated that exposure to oxygen produces severe arterial and arteriolar vasoconstriction. Wittner (38) felt that these changes lead to pulmonary hypertension. Resistance to blood flow in the pulmonary circulation of dogs has been shown to increase when their lungs are ventilated with 100% oxygen and perfused with blood recirculated only through the pulmonary circulation (39). This increase in pulmonary vascular resistance occurred when nerves were cut or blocked or abolished by inhaled isoproterenol. Elevation of intra-alveolar P_{O_2} without increase in pulmonary arterial P_{O_2} was sufficient to increase pulmonary vascular resistance. Cropp (39) felt that the pulmonary venules or veins were the probable site of the constriction. In Kistler's morphometric studies (40), the capillary volume dropped during the second and third days to half its original value. The decrease in carbon monoxide diffusing capacity during oxygen exposure which has been reported by several authors may well indicate a decrease in capillary surface area at 48-hr exposure (41, 42). Barber (43), in his study of humans with irreversible brain damage, showed that after 48 hr the oxygen-ventilated group had significantly greater dead space to tidal volume ratios than the control groups. In our laboratory, studies on capillary filtration of the lung following exposure to 100% oxygen have shown that initially there is a decrease in capillary filtration and an increase in ventilatory dead space during this time, which is indicative of the decrease in perfused capillary surface area. However, after 48 hr of exposure, dead space measurements returned to normal, and there is a marked increase in leakage of fluid across the capillary into the interstitial space as manifested by a marked increase in lung weight (44). Therefore, it appears that the initial response of the pulmonary vascular system to high oxygen tension is a diffuse pulmonary vasoconstriction. This may well be a compensatory mechanism which, as exposure time is increased, is abolished exposing a permeable, oxygen-poisoned, capillary membrane (45).

A similarity between the edema of acute pulmonary oxygen toxicity and that of left ventricular failure has raised the question of myocardial involvement. There have been only a very few studies in which the cardiovascular variables have been followed for prolonged periods during the development of pulmonary oxygen toxicity, and no definitive statements can be made at this time.

NEUROENDOCRINE SYSTEM

Animals exposed to high concentrations of oxygen underventilate, have subnormal CO_2 production, and generally appear to be at a lower metabolic rate prior to dying from oxygen poisoning. The thyroid gland appears to affect the rate of development of pulmonary oxygen toxicity. Thyroid extract was shown to hasten the onset of both convulsions and pulmonary damage in cats at hyperbaric pressures of oxygen. Even at pressure of 1 atm of oxygen, Thyroxine augments the onset of toxicity in both normal and hypophysectomized rats. Conversely, the depression of cellular metabolism by anesthesia or hibernation is associated with reduction in susceptibility to pulmonary oxygen toxicity (46).

Adrenal cortical hormones are also known to be implicated in development of oxygen toxicity. At normal pressures, prolongation in the development of pulmonary damage is afforded by adrenalectomy in rats exposed to 100% oxygen. Conversely, adrenal corticoids augment the rate of development of pulmonary oxygen toxicity (47). There is little information about the importance of the sympatho-adrenal pathways and the development of oxygen toxicity at normal pressures. Although exogenous epinephrine has been shown to augment the rate of pulmonary oxygen toxicity in rats, it is known that epinephrine toxicity alone is associated with pulmonary edema (48). Many general anesthetic agents appear to offer partial protection from pulmonary and central nervous system toxicity. Pretreatment of the animal with adrenergic blocking has been reported to provide some pulmonary protection (7).

In summary, from the evidence available, it appears that the neural endocrine system may play a role of great importance in response to hyperbaric and normabaric oxygen. Bean (10) has suggested that the variations of the neuroendocrine activity may account for the varying responses to oxygen at high pressure. There is a need for further studies in this field.

HYPERBARIC OXYGEN

Oxygen toxicity which occurs at oxygen tensions greater than one atmosphere is unique, and differs from that seen in patients exposed to high oxygen tensions at atmospheric pressures for prolonged periods. Bean (10) reviewed the effect of oxygen at hyperbaric pressure in 1945. Early signs of oxygen poisoning at hyperbaric pressures are muscular twitching, nausea, vomiting, dizziness, tunnel vision, hearing difficulty, and difficulty with breathing described as "air hunger." These symptoms are accompanied by anxiety, fatigue, and poor coordination. Grand mal convulsions of short duration are brought on by exposure to high pressure oxygen inhalation. Toxicity of this type is seen during exposure to 100% oxygen during treatment of decompression sickness.

It was formerly believed that central nervous system effects of oxygen would herald pulmonary toxicity. Bean (10) felt that the pulmonary effects of high oxygen tension were related to the CNS effect and convulsions, thus postulating a noxious extrapulmonary agent. Reich (49) pointed out that pulmonary oxygen toxicity could result rapidly with systemic hyperoxia and that convulsions (CNS toxicity) were not necessary for significant pulmonary injury.

The mechanism by which high systemic oxygen tension produces pulmonary injury is not clear. Perhaps systemic hyperoxia causes formation of an extrapulmonary agent toxic to the lung. However, it is possible that higher than normal oxygen tension in mixed venous blood may act directly upon the lung.

An extrapulmonary noxious agent was postulated by Kistler (11). Because of the chronological finding of endothelial injury before alveolar epithelial injury, he felt that a blood borne agent might be necessary for oxygen poisoning of the lung. However, the major body of evidence still points to the importance of the local toxic effect of high concentration of oxygen on the airway and lung tissue.

OXYGEN POISONING IN MAN

Almost all studies in humans are retrospective, and obvious difficulties in interpretation are present in all instances. The reason for the difficult interpretation of the total impact of oxygen poisoning to the lung is the presence of pulmonary disease prior to the treatment with high concentrations of inspired oxygen. Nash (36) felt that he could correlate greater histologic change in patients undergoing prolonged treatment with high oxygen concentrations. He distinguished two changes, an early exudative phase characterized by congestion, alveolar edema, intro-alveolar hemorrhage, and hyaline membranes, and a late proliferative phase consisting of alveolar and interlobular septal edema, alveolar hyperplasia and fibroblastic proliferation with early fibrosis. When pathologic changes in the human were compared with primate studies, it appears that 40–50 hr of exposure represents the earliest stage in which histologic changes due to oxygen toxicity become prominent. Electron microscopy might be useful in distinguishing between the patients in whom pulmonary disease may be confused with pulmonary oxygen poisoning.

Many of the physiologic responses seen in the cardiovascular system after introduction of 100% oxygen are well documented (50, 51). There is an increase in systemic vascular resistance which occurs in response to hyperoxia with decline in cardiac output. Hyperoxia has been shown to produce an increase in pulmonary arterial pressure. However, the changes in pulmonary artery pressure in dogs appear to be relatively insignificant.

A decline in dynamic pulmonary compliance in man in the absence of chest film changes has been reported (52). Atelectasis and edema have been ruled out with the absence of radiologic evidence or alteration of the A-aD 0_2. Exposure of human volunteers to high oxygen pressures has permitted measurement of altered respiratory function many hours before development of anatomical changes of irreversible toxicity (53). The rate of development of severe symptoms is variable. One of the earliest signs of pulmonary oxygen toxicity in man is chest pain. One of the major changes occurring in lung function is a change in vital capacity (41, 42). In several studies, vital capacity continued to fall after chest pain had disappeared, suggesting that the subject's discomfort was not the primary cause of such a change. Caldwell's study (41) showed that the vital capacity of one subject decreased 35% after 74 hr of exposure and that several weeks elapsed before the pre-exposure value was restored. Also, in Caldwell's subjects, the cause of impairment of vital capacity was not established, but on the basis of associated decrease in carbon monoxide diffusing capacity, beginning at 48 Hr, it was proposed that early alveolar edema formation had occurred. However the absence of radiologic changes and a low A-aD O_2 make this explanation unlikely. Clark & Lambertsen (54) suggested that the diminution of vital capacity is the best index of the development of oxygen toxicity in man.

Absorptive atelectasis has been implicated as the mechanism whereby the decrease in vital capacity is induced in human toxicity (55). However, this cannot be verified on the basis of chest film. Another possible mechanism could be muscle weakness due to fatigue and discomfort, described as retrosternal soreness induced by deep inspiration. This appears to be substantiated by the observation that reduction of vital capacity seems to occur mainly at the expense of the inspiratory reserve volume.

Other changes observed in man include abnormalities of compliance, diffusion capacity, and pulmonary blood volume, along with airway resistance. All of these studies have inherent difficulties when obtained from patients breathing 100% oxygen, and technical difficulties arise in obtaining pulmonary diffusion capacity

(DLCO) measurements or changes in compliance which might be observed but are reversed by deep breathing. This difficulty in measurement of pulmonary capillary volumes makes the significance of measured changes unclear. Studies in humans breathing 95% oxygen show a drop in diffusing capacity to 81% of control after 48 hr and to 73% after 74 hr. Total lung capacity fell to 73% of control value after 74 hr. These results might be explained on the basis of increase in blood tissue barrier, decrease in pulmonary capillary surface area, or alveolar edema formation as observed in animal studies (41, 42). Morphologic changes in man are similar to those seen in monkeys exposed to prolonged hyperoxia. Also in these studies there was a progressive fall in the diffusing capacity in patients breathing 60–100% oxygen for periods longer than three days (56).

PREVENTION AND MANAGEMENT

Pulmonary oxygen toxicity is essentially the only major manifestation of oxygen overdosage which is observed in clinical practice, with the exception of retrolental fibroplasia in neonates. Central nervous system toxicity which consists mainly of grand mal seizures is a matter of concern only in unusual environments in which oxygen at pressures greater than 2.5 atm is inspired.

High concentrations of oxygen administered for long periods of time at atmospheric pressure will cause pulmonary oxygen toxicity, but clinically it is well to remember the relative dangers of hypoxia and hyperoxia. Only in the past decade has the pathophysiology of hypoxia in the critically ill patient been commonly understood and has appropriate therapeutic manipulation become a normal part of medical practice. The knowledge that oxygen toxicity is an iatrogenic disease should be of distinct advantage in protecting patients from inappropriate use of oxygen. However, patients should not be exposed to dangerous levels of hypoxia from the fear of oxygen toxicity. Hypoxia is common and the damage caused is rapid and often irreversible.

Once oxygen therapy has been selected for a patient, there are certain guidelines that promote safe usage. Currently, it appears that pulmonary oxygen toxicity will not develop in humans if the inspired tensions are below ½ atm and it is doubtful that significant oxygen toxicity develops in patients breathing pure oxygen for less than 24 hr (57). It is evident that oxygen administered at atmospheric pressures causes damage only after prolonged exposure (55, 56, 58). There are no clinical or experimental data to indicate that patients with pre-existing pulmonary disease are more sensitive to oxygen toxicity than normal volunteers. Oxygen toxicity should not be a consideration during surgery, resuscitation, or during short term transport to a treatment center. There is no known contraindication to use of pure oxygen for brief periods in emergency situations. One must remember that the rate of development of oxygen toxicity is a function of both inspired tension and duration of exposure. Therefore, one should lower the oxygen tension as it proves feasible, but only by using objective criteria and measurement of blood gases.

Oxygen toxicity becomes a significant clinical problem with the treatment of severe, prolonged respiratory failure and respiratory insufficiency, rather than with the management of chronic lung disease, infection, or trauma. Sufficient oxygen must be administered with manipulations other than increase in inspiratory oxygen tensions in each of these conditions. Optimal ventilatory patterns, fluid restriction, maintenance of electrolyte balance, and institution of continuous positive airway pressure are all measures that may serve to increase arterial oxygen tension at a given inspired concentration (59–61). On patients requiring ventilator support, it is absolutely vital that the physician be continually aware of the patient's optimal tidal volume and respiratory rate. High pressure inspiratory values with overinflation should be avoided. Conversely, adequate inspiratory pressures and flow rate should be maintained as well as adequate sigh volumes. Parameters during ventilatory therapy should be constantly evaluated and titrated to the patient's disease process. In the clinical management of these patients, it is important to realize that as the shunt fraction increases, the increase in arterial tension for a given increment of alveolar tension is much less. As a consequence, the increase in systemic oxygenation becomes negligible at that same point. In such a patient, an increasing inspired oxygen fraction may have little benefit and the risk ot toxicity is disproportionately larger. Conversely, however, even a small reduction in shunt fraction by maneuvers such as those mentioned will accomplish far more in removing the patient from the thin line between hypoxia and oxygen toxicity. Thus, although FIO_2 is an important determinant of arterial tension, it is not the only one, and it is important to know when other manipulations may prove beneficial in maintaining the patient.

In the presence of oxygen toxicity, the physician is dealing with a more permeable pulmonary capillary membrane than is normally present. Plasma colloid osmotic pressure should therefore be maintained above 20 mm Hg by the infusion of colloid before the capillary transudation occurs. Once there is evidence that active transudation of fluid has occurred into the pulmonary parenchyma, it is exceedingly difficult to reverse this process by increasing colloid osmotic pressure. Prevention is the key in the treatment of pathophysiology of oxygen toxicity. Careful measurements of fluid balance must be maintained in order to avoid overtransfusion and daily body weight should be obtained.

Daily sputum cultures or tracheal aspirates should be obtained and chest X rays followed at regular intervals. Antibiotics should not be used unless clinical signs of pneumonia are present. The antibiotics then used should be of narrow spectrum based upon culture and gram stain results of the sputum. Steroids seem contraindicated in the therapy of pulmonary oxygen toxicity if work indicating a decrease in tolerance to oxygen when dogs are treated with steroids is applicable (47).

The treatment of acute respiratory failure is quantitative (62). Therapy is based on titration rather than prescription. The reality of oxygen toxicity is one of the factors necessitating quantification. In spite of extensive research, pulmonary oxygen toxicity remains a potentially lethal phenomenon and to date no therapeutic measures exist to influence the progressive oxygen damage except avoidance of dangerously high levels of inspired oxygen.

Literature Cited

1. Priestly, J. 1775. Reprinted 1906 in *Alembic Club Reprints,* Part I, No. 7. Chicago: Univ. Chicago Press
2. Bert, P. 1878 Translated 1943 in *Barometric Pressure: Researches in Experimental Physiology,* p. 5. Transl. M. A. Hitchcock, F. A. Hitchcock. Columbus, Ohio: Longs College Book Co.
3. Smith, J. L. 1897–98. *J. Physiol. (London)* 22:307–18
4. Smith, J. L. 1899. *J. Physiol. (London)* 24:19–35
5. Cusick, P. L., Benson, O. O. Jr., Boothby, W. M. 1940. *Proc. Mayo Clinic* 15:500
6. Gerschman, R., Nadig, P. W., Snell, A. C. Jr., Nye, S. W. 1954. *Am. J. Physiol.* 179:115–18
7. Clark, J. M., Lambertsen, C. J. 1971. *Pharmacol Rev.* 23:37–133
8. Winter, P. M., Smith, G. 1972. *Anesthesiology* 37:210–41
9. Haugaard, N. 1968. *Physiol. Rev.* 48:311–73
10. Bean, J. W. 1945. *Physiol. Rev.* 23:1–147
11. Kistler, G. S., Caldwell, P. R. B., Weibel, E. R. 1967. *J. Cell Biol.* 32:605–28
12. Kaplan, H. P., Robinson, F. R., Kapanci, Y., Weibel, E. R. 1969. *Lab. Invest.* 20:94–100
13. Kapanci, Y., Weibel, E. R., Kaplan, H. P., Robinson, F. R. 1969. *Lab. Invest.* 20:101–18
14. Pratt, P. C. 1965. *Ann. N. Y. Acad. Sci.* 121:809–22
15. Wolfe, W. G., Sabiston, D. C. Jr. 1973. *Surg. Gynecol. Obstet.* 137:763–68
16. Penrod, K. E. 1958. *Fed. Proc.* 17:123
17. Wolfe, W. G., Ebert, P. A., Sabiston, D. C. Jr. 1972. *Surgery* 72:246–52
18. Miller, W. W. et al 1970. *N. Engl. J. Med.* 282:943–47
19. Ashbaugh, D. G. 1971. *J. Appl. Physiol.* 31:664–68
20. Thomas, A. N., Ketchm, S. A. III, Hall, A. D., Zubrin, J. R. 1970. *Proc. 4th Int. Congr. Hypb. Med.,* pp. 22–28
21. Brauer, R. W., Parrish, D. C., Way, R. O., Pratt, P. C., Pessotti, R. L. 1970. *J. Appl. Physiol.* 28:474–81
22. Davies, H. C., and Davies, R. E. 1965. Biochemical aspects of oxygen poisoning. *Handbook of Physiology* 11(3)-:1047–58
23. Chance, B., Jamieson, D., Coles, H. 1965. *Nature* 206:257–63
24. Dickens, F. 1946. *Biochm J.* 40:145–87
25. Sanders, A. P., Hall, I. H., Woodhall, B. 1965. *Science* 140:1830–31
26. Fridovich, I. 1972. *Accounts Chem. Res.* 5:321–26
27. McCord, J. M., Beauchamp, C. O., Goscin, S., Misra, H. P., Fridovich, I. 1973. *Proc. 2nd Int. Symp. Univ. Park Press, Baltimore,* pp. 51–76
28. McCord, J. M., Fridovich, I. 1969. *J. Biol. Chem.* 244:6049–55
29. Clarke, G. M., Smith G., Sandison, A. T., Ledingham, I. M. 1973. *Am. J. Physiol.* 224:248–55
30. Edmunds, L. H. Jr., Holm, J. C. 1968. *J. Appl. Physiol.* 24(2):115–23
31. Morgan, T. E., Finley, T. N., Huber, G. L., Fialkow, H. 1965. *J. Clin. Invest.* 44:1737–44
32. Lee, C. J., Lyons, J. H., Konisberg, S., Morgan, F., Moore, F. D. 1967. *J. Thorac. Cardiov. Surg.* 53:759–69
33. Motlagh, F. A. et al 1969. *Surg. Forum* 20:219–20
34. Coalson, J. J., Beller J. J., Greenfield, L. F. 1971. *J. Pathol.* 104:267–73
35. Nash, G., Bowen, J. A., Langlinais, P. C. 1971. *Arch. Pathol.* 91:234:240
36. Nash, G., Blennerhassett, J. B., Pontoppidan, H. 1967. *N. Engl. J. Med.* 276:368–74
37. Laurenzi, G. A., Yin, S., Guarneri, J. J. 1968. *N. Engl. J. Med.* 279:333–39
38. Wittner, M., Rosenbaum, R. M. 1966. Pathophysiology of pulmonary oxygen toxicity. *Proc. 3rd Int. Conf. Hypb. Med.,* ed. I. W. Brown Jr., B. G. Cox, 179–88. Washington, DC: Nat. Acad. Sci.
39. Cropp, G. J. A. 1965. *Am. J. Physiol.* 208:130–38
40. Kistler, G. S., Caldwell, P. R. B., Weibel, E. R. 1966. Quantitative electron microscopic studies of murine lung damage after exposure to 98.5% oxygen at ambient pressure: A preliminary report. See ref. 38, pp. 169–78
41. Caldwell, P. R. B., Lee, W. L. Jr., Schildkraut, H. S., Archibald, E. R. 1966. *J. Appl. Physiol.* 21:1477–83
42. Fisher, A. B., Hyde, R. W., Puy, R. J. M., Clark, J. M., Lambertsen, C. J. 1968. *J. Appl. Physiol.* 24:529–35
43. Barber, R. E., Lee, J., Hamilton, W. N. 1970. *N. Engl. J. Med.* 282:1478–84
44. Wolfe, W. G., DeVries, W. C., Anderson R. W., Sabiston, D. C. Jr. 1974. *J. Surg. Res.* 16:312–17
45. DeVries, W. C., Anderson R. W., Wolfe

W. G., Alexander, L. G. 1973. *Surg. Forum* 24:224–26

46. Smith, C. W., Bean, J. W., Bauer, R. 1960. *Am. J. Physiol.* 199:883–89
47. Gerschman, R., Culbert, D. L., Nye, S. W., Nadig, P. W., and Fenn, W. O. 1954. *Am. J. Physiol.* 178:346–50
48. Ersoz, N., Finestine, S. C. 1971. *Brief Anaesth.* 43:709–12
49. Reich, T., Tait, J., Suga, T., Naftchi, N. E., Demeny, M., 1972. *J. Appl. Physiol.* 32:374–79
50. Bergofsky, E. H., Bertun, P. 1966. *J. Appl. Physiol.* 21:567–72
51. Eggers, G. W. N. Jr., Paley, H. W., Leonard, J. J., Warren, J. V. 1962. *J. Appl. Physiol.* 17:75–79
52. Fisher, A. B., Hyde, R. W., Puy, R. J. M., Clark, J. M., Lambertsen, C. J. 1968. *J. Appl. Physiol.* 24:529–36
53. Burger, E. J. Jr., Mead, J. 1969. *J. Appl. Physiol.* 27:191–97
54. Clark, J. M., Lambertsen, C. J. 1971. *J. Appl. Physiol.* 30:739–52
55. Michel, E. L., Langevin, R. W., Gell, C. F., 1960. *Aerosp. Medicine* 31:138–44
56. Kapanci, Y., Tosco, R., Eggermann, J., Gould, V. 1972. *Chest* 62:162–69
57. VanDeWater, J. M. et al 1970. *N. Engl. J. Med.* 283:621–26
58. Singer, M. M., Wright, F., Stanley, L. K., Roe, B., Hamilton, W. K. 1970. *N. Engl. J. Med.* 283:1473–78
59. Barach, A. L., Segal, M. S. 1972. *Ann. Allergy* 30:113–21
60. Petty, T. L., Stanford, R. E., Neff, T. A. 1971. *Ann Intern. Med.* 75:361–67
61. Ashbaugh, D. G., Petty, T. C. 1973 *J. Thorac. Cardiov. Surg.* 65:165–70
62. Pontoppidan, H., Geffin, B., Lowenstein, E. 1972. *N. Engl. J. Med.* 287:690–98, 743–52, 799–806

Copyright 1975. All rights reserved

MECHANISMS OF SUPRAVENTRICULAR TACHYCARDIAS

❖7144

Bruce N. Goldreyer, M.D.
Department of Medicine, University of Southern California, Los Angeles, California 90033

BACKGROUND

Rational therapy of disease is predicated on the notion that understanding disease mechanisms will allow specific therapeutic intervention. This has clearly been the case in such disciplines as infectious disease and immunology, but the therapy of cardiac arrhythmias has been empiric at best. With the development of techniques by which the cellular and tissue electrophysiology of heart muscle could be studied, and, more recently, with the application of intracardiac electrocardiography to clinical arrhythmia analysis, the mechanisms by which these arrhythmias are initiated and sustained are gradually being elucidated (1). In supraventricular tachycardias more than any other arrhythmia, these discoveries have drastically altered clinical therapeutics.

The two basic mechanisms responsible for all cardiac arrhythmias are automaticity and reentry. Automaticity initiates cardiac arrhythmias when cells other than those in the sinoatrial node undergo spontaneous diastolic depolarization, eventually reaching threshold potential and causing a propagated premature depolarization. When the rate of diastolic depolarization exceeds that of the sinus node, repetitive firing results in an ectopic tachycardia.

In the case of reentry, abnormal conduction, rather than automaticity, is the cause of the abnormal rhythm. Prolonged conduction of a propagating wave front, between or within cardiac chambers, results in the recovery of previously inexcitable tissue, and the return of the impulse to its point of origin. If the relationship between conduction delay and tissue refractoriness is appropriate, a continuous "circuiting" of electrical activity results in a sustained tachyarrhythmia in the absence of ectopic "pacemakers."

The controversy over the mechanism of classical paroxysmal supraventricular tachycardia (SVT) [synonyms: paroxysmal atrial tachycardia (PAT); paroxysmal nodal (or junctional) tachycardia; reentrant tachycardia; reciprocal tachycardia] has recently been settled. Several independently working investigators have demonstrated, without doubt, that classical SVT is a reentrant arrhythmia. The evidence for this conclusion and its therapeutic consequences form the subject of this discussion.

EVIDENCE FOR THE REENTRANT NATURE OF PAROXYSMAL SVT

Atrial Premature Depolarizations and SVT

Although the relationship between SVT and single atrial premature depolarizations (APDs) has been recognized since the earliest days of clinical electrocardiography, it was originally thought that this relationship reflected ectopic automatic cells being occasionally capable of firing repetitively, thereby resulting in a sustained tachycardia (2). Recently, however, it was noted that APDs occurring late in the atrial cycle (i.e. those which were not very premature) never resulted in sustained tachycardias, whereas those which occurred earlier almost always did (3). The obvious difference was that APDs occurring earlier in the atrial cycle entered the atrioventricular (A-V) node earlier in its relative refractory period (that portion of the cycle during which impulses are conducted with progressive conduction delay) and uniformly demonstrated considerable delay in A-V conduction prior to the initiation of SVT. This A-V conduction delay allowed the propagating impulse to return to the chamber of its origin (the atrium). This single reentry circuit (termed an atrial "echo") results in SVT when the A-V node has recovered sufficiently after atrial reentry to once again begin antegrade conduction (4). Thus, patients with APDs and SVT characteristically demonstrate single and multiple atrial echo beats (3). When the refractoriness of the A-V conduction system and atrium are appropriate, sufficient antegrade conduction delay of the initiating impulse results in self-perpetuating reentry: SVT(5).

Atrial Stimulation and SVT

If the essential condition for SVT initiation is prolonged A-V conduction, any entry which results in sufficient prolongation of antegrade conduction in susceptible patients should result in sustained SVT. Thus, not just spontaneous, but stimulated, atrial premature beats should result in SVT, provided conduction is sufficiently prolonged (6). In more than forty patients studied by the author, single stimulated atrial premature depolarizations uniformly result in SVT when they are evoked with sufficient prematurity (Figure 1).

These single stimulated APDs are evoked well outside of the atrial vulnerable period, but even if reentry via the A-V conduction system were proven, proof of the primacy of A-V conduction delay over initiating impulse prematurity would still be necessary. A physiologic mechanism by which A-V conduction could be prolonged independently of premature atrial stimulation was found by pacing the atrium at fixed rates in excess of that of the sinus node. In sinus tachycardia the same influences which speed the sinus rate shorten A-V conduction, but during fixed rate

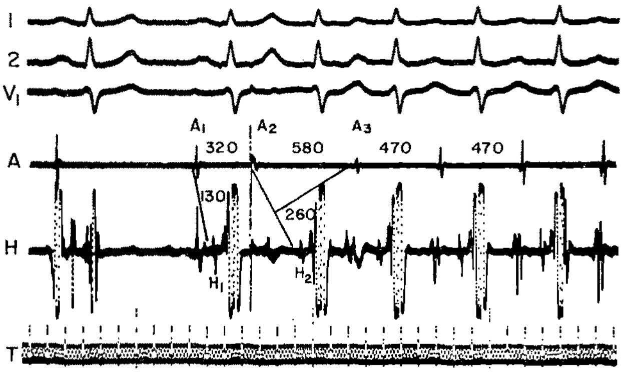


Figure 1 Onset SVT. Recorded are: ECG leads 1, 2, V_1, atrial electrogram (*A*), His bundle electrogram (*H*), and time marks at 10- and 100-msec intervals (*T*). A sinus beat (A_1) demonstrates normal A-V nodal conduction (A_1H_1=130 msec). A stimulated atrial premature depolarization (A_2), with markedly prolonged A-V nodal conduction (A_2H_2=260 msec), reenters the atrium (A_3) and begins an episode of SVT at a cycle length of 470 msec.

atrial pacing, the non-rate-dependent A-V refractoriness causes progressive A-V conduction delay at increasing paced rates. Thus, A-V conduction could be prolonged independently of premature atrial stimulation by pacing the atrium at increasing rates. In patients with paroxysmal SVT, as the paced rate is increased, SVT is uniformly produced (7). Although the cycle length of the atrial pacing is considerably longer than the coupling interval of single APDs necessary to produce SVT, once sufficient prolongation of A-V conduction occurs, self-sustaining SVT is initiated. This simple clinical electrophysiologic experiment effectively demonstrates that A-V conduction delay is the essential condition for the initiation of SVT.

AV Conduction and SVT

If SVT is a function of A-V conduction and refractoriness, then the physiological properties of the A-V conduction system will determine the electrophysiologic characteristics of SVT (3). Two simple observations prove this. First, when SVT is initiated, the shorter the coupling interval of the initiating beat, the longer the interval before the first beat of SVT. This statistically significant relationship stems from the fact that using the A-V conduction system for reentry, the earlier the initiating beat occurs, the more slowly it is conducted. And the more slowly it is conducted, the later the first beat of SVT will appear.

There is also a linear relationship between A-V conduction system refractoriness and the rate of SVT. Patients with slowly conducting A-V nodes have SVT at slower rates than those patients in whom A-V refractoriness is less prolonged. The relationship between A-V refractoriness and the rate of SVT is also statistically significant from patient to patient (3). Moreover, in the same patient, pharmacologic maneuvers designed to alter A-V conduction uniformly changed the rate of subsequent SVT. For example, after atropine administration in amounts sufficient to shorten the refractoriness of the A-V conduction system, the rate of evoked SVT is always more rapid.

The Primacy of AV Conduction

Although clearly implicating the A-V conduction system as the site, and A-V conduction system delay as the mechanism of SVT, all previous observations showed atrial premature depolarizations, either spontaneous or evoked, to be the primary event in initiated SVT. Two experiments of nature were required to prove the primacy of A-V conduction delay. In these two unique patients, A-V dissociation [in one case due to sinus slowing and emergence of a His bundle pacemaker (8); in the other, to two ventricular premature beats in a row (9)] resulted in markedly prolonged A-V conduction of a normally occurring *sinus* impulse. In both patients, this impulse reentered within the A-V conduction system and resulted in sustained SVT. Thus, even an impulse of sinus node origin may result in SVT, provided A-V conduction delay is sufficient.

The Sequence of Atrial Depolarization

Although the relationship between A-V conduction system refractoriness and the characteristics of SVT provide the basic evidence that reentry is occurring within the A-V conduction system, further evidence is provided by the atrial depolarization sequence during SVT. During sinus rhythm, atrial depolarization normally proceeds from the high right atrium to the region of the low right atrium and A-V node (Figure 2*A*). During SVT, however, since each atrial depolarization is initiated via reentry eventually terminating in A-V nodal activation, the atrial activation sequence ought to be from low to high right atrium. This is clear in all cases of SVT where P waves are distinguishable from QRS complexes (Figure 2*B*), but may also be demonstrated under unusual circumstances, even when P waves and QRSs are simultaneous (7). Thus, the inverted P waves of "reciprocating tachycardia" actually occur in all cases of SVT. They are, however, inapparent from the surface ECG when obscured by QRS or T waves.

THE SITE OF REENTRY IN SVT

The AV Node

Electrophysiological evidence has clearly demonstrated that most cases of SVT result from reentry within the A-V node. In all the preceding arguments one could substitute the phrase "A-V node" for "A-V conduction system" and almost always be accurate. Since, as stated, conduction delay is the essential condition of reentry, Figure 1 shows what must be the reentry location. It is the atrium-His (A-H) interval, or A-V nodal conduction time, of the premature depolarization which *alone* is prolonged. The A-V node, therefore, must be the location of reentry (10).

This simple, telling, and yet indirect evidence that the A-V node was the site of reentry in SVT was subsequently confirmed by microelectrode techniques in isolated tissue preparations (11, 12). In these studies, reentry could not only be localized to the A-V node, but participation of the His bundle and/or sinus node could be excluded by their surgical extirpation (12). The necessity of intact A-V nodal conduction was similarly demonstrated when mechanical A-V node stimulation in these preparations precluded SVTs initiation.

Recent evidence suggests that in some patients with SVT, the A-V node is physiologically dissociated into two functional pathways (13, 14). Classically designated "alpha" and "beta," the pathway with slow conduction and yet a short refractory period is responsible for the antegrade propagation of the impulse, whereas the return pathway has more rapid conduction but a longer refractory period (15). These dual pathways have been demonstrated in several patients with SVT, but their existence is not *required* for A-V nodal reentry. Reentry may occur even in linear strands of tissue, provided conduction is slow enough.

Other Sites of Reentry

Reentrant SVT using the distal A-V conduction system has recently been described (16, 17). In these rare cases, A-V nodal conduction of the impulses initiating tachycardia have been normal, but infranodal conduction has been prolonged. Conduction occurs via one fascicle (the right bundle for example) while the other is physiologically refractory. Ventricular activation occurs, during which time the "blocked" fascicle recovers excitability and returns the impulse to the common bundle, A-V node, and atrium. If conduction times and refractoriness are precisely

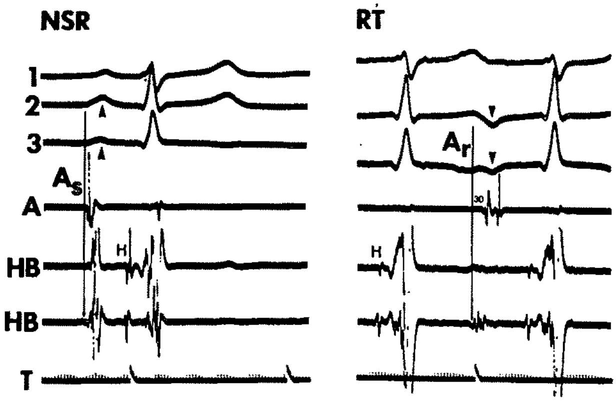


Figure 2 Atrial activation. Recorded are: ECG leads 1, 2, 3, and high right atrial electrogram (*A*), two His bundle electrograms (*H*), and time marks at 10-msec intervals. During sinus rhythm (*NSR*), atrial activation occurs first in the high right atrium. During supraventricular tachycardia (*RT*), atrial activation begins in the low right atrium adjacent to the A-V node. Thirty msec are required before the impulse conducts to the high right atrium. Arrows demonstrate inverted P waves in the inferiorly directed ECG leads.

matched, sustained reentrant tachycardia may occur. Such patients are extremely rare.

Since conduction within the sinus node is physiologically extremely slow, this location too may serve as a site of reentry and result in a sustained tachycardia (18). Rare cases have recently been reported, but they should not be confused with paroxysmal SVT (19). In these patients, the much slower sinoatrial conduction results in much slower rates of tachycardia, and most importantly, the sequence of atrial activation will always be from high to low right atrium (20).

Extranodal pathways may also sustain atrial reentry and paroxysms of tachycardia. A voluminous literature now exists in which participation of extranodal A-V connections are responsible for SVT in patients with the Wolff-Parkinson-White syndrome (21). These patients may be easily distinguished from those with classical SVT, however, since by atrial pacing or premature stimulation the characteristic delta wave can be produced even if it is not apparent during sinus rhythm. The delta wave resulting from ventricular preexcitation clearly demonstrates the existence of two functionally distinct pathways.

Although A-V conduction is characteristically short in patients with accelerated A-V conduction of the Lown-Ganong-Levine type, most evidence nonetheless supports the A-V node as the usual site of reentry responsible for their frequent bouts of paroxysmal tachycardia (22).

THE THERAPY OF PAROXYSMAL SVT

If the mechanism of paroxysmal supraventricular tachycardia has been accurately established, then effective therapies are clear. Since the vast majority of SVTs are initiated and sustained by atrial reentry via the A-V node, any maneuver which alters A-V nodal conduction and refractoriness should affect SVT. More specifically, interventions rendering the A-V node refractory to conduction should result in an inability to sustain SVT.

Vagal Maneuvers

The effect of enhanced cardiac vagal discharge is the increased release of acetylcholine at cardiac receptor sites. The A-V node has rich parasympathetic innervation, and the net effect of increased acetylcholine release at the A-V node is to prolong its conduction and refractoriness. It is not surprising therefore that therapies designed to enhance vagal tone are effective in terminating episodes of SVT (10). As demonstrated in Figure 3, the Valsalva's maneuver progressively prolongs A-V nodal conduction from 230 to 300 msec before a final atrial reentrant beat finds the A-V node so refractory that it can no longer conduct. As soon as A-V nodal conduction is interrupted, reentry is impossible and SVT terminates. Carotid sinus massage and other vagotonic maneuvers no longer considered necessary (eye ball pressure, gagging, and rectal stimulation), all terminate SVT by an identical mechanism.

Though seemingly paradoxical, the effect of pressor drugs such as metaraminol and phenylephrine is identical, causing peripheral arterial constriction and an increase in systemic arterial pressure. Aortic arch and carotid sinus barroreceptors

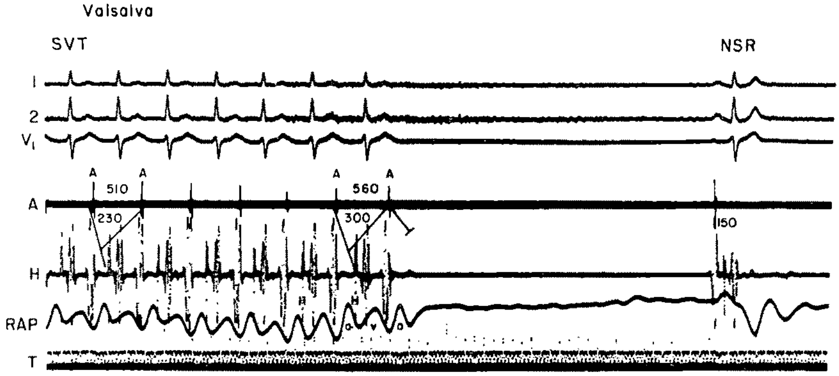


Figure 3 SVT is terminated by the Valsalva's maneuver. Abbreviations are as before, but in addition right atrial pressure (*RAP*) is recorded, and characteristic A and V waves are shown. During SVT at cycle length of 510 msec (A–A = 510 msec), A-V nodal conduction requires 230 msec. With Valsalva's maneuver, A-V nodal conduction prolongs to 300 msec and a final atrial echo fails to conduct thus terminating SVT. During sinus rhythm (*NSR*) an A-V conduction time of 150 msec is shown.

sense this pressure elevation and vagal efferent discharge is enhanced. Again, A-V nodal conduction delay terminates SVT. The 50-msec increment in the cycle length of SVT prior to its termination demonstrated in Figure 3 is the rule. Despite myriad textbook claims to the contrary, all therapies which terminate SVT by increasing A-V nodal conduction will slow the rate of SVT prior to its termination. Again, since cycle length is a function of conduction time through the A-V node, this is a necessary event observed in all vagal maneuvers which terminate SVT.

Drug Therapy

To be effective in the prophylaxis of SVT, a drug must prevent the electrical events which predispose to the A-V conduction delay required to initiate SVT (APDs for example), and/or so alter A-V nodal conduction that reentry is no longer a physiologic possibility. Seemingly antithetical drugs such as digitalis glycosides, quinidine, procainamide, and propranolol are all effective in terminating episodes of SVT and preventing their recurrence because they prolong A-V nodal conduction either directly or indirectly, and the delicate balance between A-V nodal refractoriness and A-V nodal conduction is upset (23). If that balance is sufficiently altered, SVT becomes a physiological impossibility (6).

The ability to initiate SVT at will by premature atrial stimulation has allowed immediate drug efficacy evaluation and obviated the need for prolonged drug trials. In Figure 4, after the administration of 5 mg of propranolol IV, a single induced APD is still capable of initiating SVT. Contrasting Figures 1 and 4 demonstrates that propranolol is ineffective because A-V nodal refractoriness has not been altered.

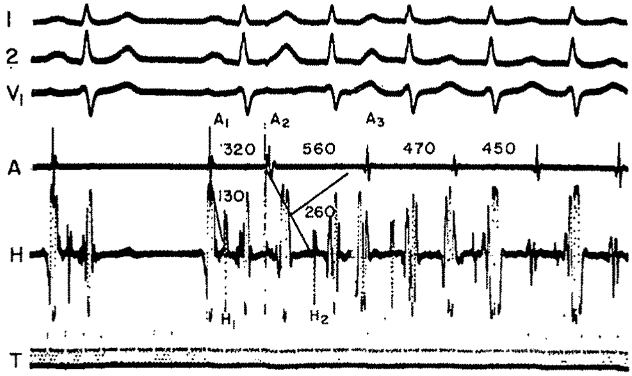


Figure 4 Identical to Figure 1 except that 5 mg of propranolol have been administered (see text for discussion).

Electrical Therapy

If SVT results from A-V nodal reentry, then premature depolarization of either atria or ventricles, rendering part of the reentrant pathway refractory, should be sufficient to terminate episodes of SVT. Just as atrial premature depolarizations initiate SVT, so stimulated atrial premature beats may terminate it. This finding, uniformly demonstrated in patients with SVT (3), and initially utilized as evidence for reentry, now forms the basis of an entire new mode of therapy (1, 24–26).

In Figure 5, during an episode of SVT at a cycle length of 470 msec, the atrium is prematurely stimulated after 320 msec. This atrial premature depolarization

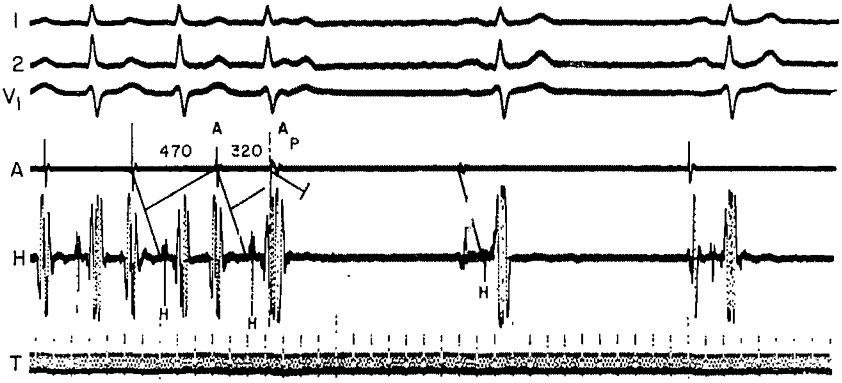


Figure 5 SVT terminated by a single stimulated atrial premature depolarization (A_p). Abbreviations as before. During SVT at a cycle length of 470 msec, a single A_p is introduced 320 msec into the atrial cycle. The A-V node is refractory and SVT terminates.

enters the A-V node, not only rendering the return pathway to the atrium refractory, but also entering the A-V node when antegrade conduction cannot be completed. A-V nodal reentry is terminated, and this episode of SVT is, naturally, terminated as well.

Similarly, premature stimulation of the ventricles also terminates SVT. The premature depolarization enters the A-V node and renders the reentrant pathway refractory to antegrade impulse propagation. Reentry thus interrupted, SVT ends (27). Atrial or ventricular permanent demand pacemakers are being implanted with increasing frequency to terminate episodes of SVT in patients refractory or intolerant to drug therapy.

CONCLUSION

The electrophysiologic mechanism of a specific cardiac arrhythmia has been elucidated by intracardiac recording and stimulating techniques. The documentation that SVT is initiated and sustained by A-V nodal reentry has explained the rationale of therapies heretofore empiric, and allowed the development of new modes of therapy for its prophylaxis and termination. Using this arrhythmia as a model, one may reasonably expect that similar investigation of other frequent and disabling cardiac arrhythmias will result in rational therapeutic interventions.

Literature Cited

1. Goldreyer, B. N. 1972. Intracardiac electrocardiography in the analysis and understanding of cardiac arrhythmias. *Ann. Intern. Med.* 77:117–36
2. Bellet, S. 1963. *Clinical Disorders of the Heart Beat.* Philadelphia: Lea & Febiger. 282 pp.
3. Bigger, J. T. Jr., Goldreyer, B. N. 1970. The mechanism of supraventricular tachycardia. *Circulation* 42:673–88
4. Wallace, A. G., Daggett, W. M. 1964. Re-excitation of the atrium. "The echo phenomenon." *Am. Heart J.* 68:661–66
5. Moe, G. K., Cohen, W., Vick, R. L. 1963. Experimentally induced paroxysmal A-V nodal tachycardia in the dog. *Am. Heart J.* 65:87–92
6. Goldreyer, B. N., Bigger, J. T. Jr. 1969. Spontaneous and induced reentrant tachycardia. *Ann. Intern. Med.* 70:87–98
7. Goldreyer, B. N., Damato, A. N. 1971. The essential role of atrioventricular conduction delay in the initiation of paroxysmal supraventricular tachycardia. *Circulation* 43:679–87
8. Goldreyer, B. N., Weiss, M. B., Damato, A. N. 1971. Supraventricular tachycardia initiated by sinus beats. *Circulation* 44:820–25
9. Kastor, J. A., Goldreyer, B. N. 1973. *Cardiac Arrhythmias.* New York: Grune & Stratton. 391 pp.
10. Goldreyer, B. N., Bigger, J. T. Jr. 1971. Site of reentry in paroxysmal supraventricular tachycardia in man. *Circulation* 43:15–26
11. Janse, M. J., Van Capelle, F. J. L., Freud, G. E., Durrer, D. 1971. Circus movement within the A-V node as a basis for supraventricular tachycardias as shown by multiple microelectrode recording in the isolated rabbit heart. *Circ. Res.* 28:403–14
12. Wit, A. L., Goldreyer, B. N., Damato, A. N. 1971. An in vitro model of paroxysmal supraventricular tachycardia. *Circulation* 43:862–75
13. Denes, P., Wu, D., Dhingra, R. C., Chuquima, R., Rosen, K. M. 1973. Demonstration of dual A-V nodal pathways in patients with paroxysmal supraventricular tachycardia. *Circulation* 48:549–55
14. Rosen, K. M., Mehta, A., Miller, R. A. 1974. Demonstration of dual atrioventricular nodal pathways in man. *Am. J. Cardiol.* 33:291–94
15. Mendez, C., Moe, G. K. 1966. Demonstration of a dual A-V nodal conduction

system in the isolated rabbit heart. *Circ. Res.* 19:378–94

16. Lozano, J., Mandel, W. J., Hayakawa, H., Shine, K. I., Eber, L. M. 1973. Reentrant tachycardia: participation of the distal A-V conduction system. *Chest* 63:23–26
17. Josephson, M. E., Varghese, J. P. 1974. Intraventricular reentry. *Chest* 65:237
18. Han, J., Malozzi, A. M., Moe, G. K. 1968. Sino-atrial reciprocation in the isolated rabbit heart. *Circ. Res.* 22: 355
19. Childers, R. W., Arnsdorf, M. F., de la Fuente, D. J., Gambetta, M., Svenson, R. 1973. Sinus node echoes. *Am. J. Cardiol.* 31:220
20. Narula, O. S. 1972. Sinus node re-entry: mechanism of supraventricular tachycardia (SVT) in man. *Circulation* 45–46:27 (Abstr.)
21. Boineau, J. P., Moore, E. N., Spear, J. F., Sealy, W. C. 1973. *Cardiac Arrhythmias.* New York: Grove & Stratton. 421 pp.
22. Caracta, A. R. et al 1973. Electrophysiological studies in the syndrome of short P-R interval, normal QRS complex. *Am. J. Cardiol.* 31:245
23. Gettes, L. S., Yoshonis, K. F. 1970. Rapidly recurring supraventricular tachycardia. A manifestation of reciprocating tachycardia and an indication for propranolol therapy. *Circulation* 41:689–700
24. Lister, J. W., Cohen, L. S., Bernstein, W. H., Samet, P. 1968. Treatment of supraventricular tachycardias by rapid atrial stimulation. *Circulation* 38:1044–59
25. Hunt, N. C., Cobb, F. R., Waxman, M. B. 1968. Conversion of supraventricular tachycardias with atrial stimulation. Evidence for reentry mechanisms. *Circulation* 38:1060–65
26. Schnitzler, R. N., Goldreyer, B. N., Westura, E. E., Damato, A. N. 1971. Demand coronary vein pacemakers for drug-resistant supraventricular tachycardia. *Circulation* 43–44 (Suppl. II):97 (Abstr.)
27. Kitchen, J. G. III, Goldreyer, B. N. 1972. Demand pacemaker for refractory paroxysmal supraventricular tachycardia. *N. Engl. J. Med.* 287:596–99

Copyright 1975. All rights reserved

SELECTIVE DISTAL SPLENORENAL SHUNT FOR BLEEDING ESOPHAGEAL VARICES[1]

❖7145

Robert B. Smith III, M.D. and W. Dean Warren, M.D.
The Joseph Brown Whitehead Department of Surgery, Emory University School of Medicine, Atlanta, Georgia 30322

INTRODUCTION

Gastrointestinal hemorrhage constitutes a major threat to cirrhotic patients with portal hypertension and esophageal varices. Whipple and Blakemore reintroduced the concept of portal-systemic decompression for relief of bleeding varices in the mid-1940s, and ushered in an era of portacaval shunting which has continued until the present (1, 2). Over the intervening thirty years, a number of modifications of the basic portacaval shunt operation have been introduced, including the use of prosthetic interposition grafts, but all serve to accomplish total decompression of the portal venous system into a lower pressure systemic vein, thus bypassing the liver with a major portion of the portal venous flow (3). As experience has accumulated with total shunted patients, it has become evident that certain individuals tolerate the procedure poorly and many suffer major morbidity or decreased survival when compared with nonshunted cirrhotics. Results of controlled, randomized studies from Boston and New Haven demonstrated clearly that there was no advantage to "prophylactic" portacaval shunts, as any potential increased survival resulting from prevention of variceal hemorrhage was more than offset by an increased rate of postoperative hepatic failure (4, 5). Other studies have suggested that even in patients undergoing "therapeutic" shunts, the five-year survival is not significantly different between surgical and nonsurgical populations. Only the mode of death is changed with the operative patients dying less often from massive hemorrhage (6). In addition to this rather disappointing record in respect to long-term survival,

---

[1] Supported in part by National Institutes of Health Research Grant AM 15736 and General Clinical Research Center Public Health Service Research Grant 5M01RR00039.

shunts are also responsible for severe morbidity in the form of portal-systemic encephalopathy in a high percentage of patients (7).

Dissatisfaction with the available methods of surgical treatment for portal hypertension prompted introduction of the selective distal splenorenal shunt by Warren, Zeppa & Fomon in 1967 (8). This procedure was designed to achieve three fundamental goals: (*a*) There must be selective decompression of gastroesophageal varices to assure adequate control of hemorrhage; (*b*) Portal venous perfusion of the liver should be preserved to the greatest extent possible to reduce the likelihood of hepatic failure or metabolic encephalopathy; and (*c*) High intestinal venous pressure should be maintained to assure hepatopetal flow of mesenteric vein blood and possibly to help reduce the tendency for protein intoxication, as suggested by the animal experiments of Price et al (9).

Essentials of the selective distal splenorenal shunt procedure, therefore, include the following surgical maneuvers: (*a*) Venous pressure in the spleen is reduced to near systemic pressure levels by anastomosing the distal (splenic) end of the splenic vein to the side of the left renal vein; (*b*) The short gastric veins and other collateral vessels that can decompress esophageal varices into the low-pressure splenic venous system are carefully preserved; and (*c*) The remainder of the portal circulation is isolated, insofar as possible, from the newly created gastrosplenic venous compartment by ligation of the coronary vein and the umbilical vein (Figure 1).

PATIENT SELECTION AND PREOPERATIVE MANAGEMENT

Every effort is made to evaluate and prepare each patient carefully before the surgical procedure is undertaken. Active bleeding from esophageal varices should be controlled by nonoperative means, such as gastric lavage, administration of Pitressin® or use of the Sengstaken-Blakemore tube. After the patient is stable, a complete assessment of his general condition is made, including liver function tests and liver biopsy, in order to assess suitability for operation. Individuals with tense ascites, or those in whom moderate ascites is resistant to diuretic therapy, are excluded from consideration since this shunt may result in failure to control ascites with resultant high postoperative morbidity. Some modification of the standard side-to-side type portal decompression would be necessary in such cases.

Proper timing of the operation is another important consideration in these patients. Individuals in a high-risk, or Child's class-C, category may benefit from a delay in the operation to afford time for a preoperative build-up program. Postponement of the procedure for a period of 14–21 days often results in improvement of liver functional reserve and, therefore, a lowered operative risk. Unfortunately, recurrent bleeding necessitates abandonment of the build-up program in a few patients. If necessary, the distal splenorenal shunt can be performed as an emergency procedure with the Sengstaken-Blakemore balloons inflated.

An essential part of each patient's preoperative evaluation is the performance of angiography to assure patency of the portal venous trunks and to define the anatomic relationship of the splenic vein to the left renal vein. An angiographic "liver

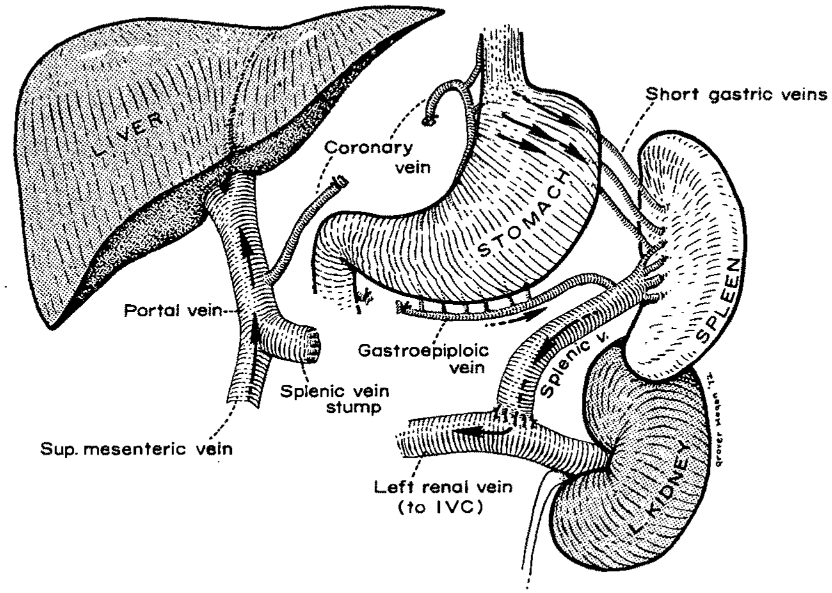


Figure 1 Semi-diagrammatic illustration of the distal splenorenal anastomosis. Note transplenic decompression of varices while maintaining portal venous perfusion of the liver.

package" is obtained on each patient, including hepatic wedge pressure and hepatic wedge venogram, left renal venogram,and superior mesenteric/splenic arteriograms followed through the venous phases. Percutaneous splenoportography is performed only if the above indirect techniques fail to provide adequate visualization, since any injury to the spleen sufficiently severe to require splenectomy obviously would preclude distal splenorenal shunt.

OPERATIVE MANAGEMENT

The peritoneal cavity is entered through a long, bilateral subcostal incision extending from the midclavicular line on the right to the anterior axillary line on the left (10). Care is taken in exploration of the abdomen to avoid injury to vascular adhesions in areas that will not be decompressed by the splenic region shunt. The lesser omental bursa is entered via the gastrocolic omentum, deliberately ligating the gastroepiploic vein at the pylorus, but carefully preserving the short gastric veins when they are encountered along the greater curvature. The posterior parietal peritoneum is then incised along the lower border of the pancreas and that gland gently reflected anteriorly and cephalad exposing the retropancreatic course of the splenic vein. The inferior mesenteric vein may be encountered and should be divided and ligated as its confluence with the splenic vein trunk. Next, the central portion of the splenic vein is freed from the pancreas by identifying and carefully suture-ligating the several venous branches which connect the two structures. This is the most tedious and potentially troublesome part of the procedure as the small pancreatic venous branches are thin-walled and bleed freely if torn. After a suitable length of splenic vein is mobilized, attention is directed to the preparation of the left renal vein. Since the renal vein lies beneath a thick pad of retroperitoneal tissue, it is quite helpful to know its approximate anatomic position before beginning to dissect posteriorly. This information is available by comparing the relative positions of the splenic and renal veins on the preoperative angiograms; palpation of the lower pole of the left kidney through the retroperitoneum also serves as a guide to its position. As the dissection is carried down to the vein wall, hemostasis can best be maintained by methodically dividing the tissue between clamps and ligating each segment. It may be necessary to divide the adrenal vein and gonadal vein in order to free a sufficient length of renal vein trunk to permit application of a partial occlusion clamp. Attention is then returned to the splenic vein, which is divided close to its junction with the superior mesenteric vein and the mesenteric stump closed with a continuous vascular suture. The site of proposed end-to-side venous anastomosis is then determined on the anterosuperior aspect of the left renal vein and an ellipse of vein wall excised. The anastomosis is performed with continuous sutures on the posterior row and interrupted sutures on the anterior row in an effort to avoid any purse-string effect on the opening. Following completion of the shunt, the operator must identify and disconnect the coronary vein as its courses through the retroperitoneum between the superior border of the pancreas and the esophagogastric junction. In closing the incision, a careful water-tight approximation of the peritoneum and posterior fascia should be made to minimize the possibility of ascitic fluid leak.

The chief disadvantage of the distal splenorenal anastomosis in comparison with various other shunting operations is acknowledged to be its technical difficulty. It should be noted, however, that with reasonable technical skill and attention to hemostasis, the procedure can be done almost routinely within three hours operating time and has been performed frequently by the authors without need for blood transfusion. If technical considerations such as tortuosity or potential kinking of the splenic vein produce concern about eventual patency of the shunt, a modified distal splenorenal decompression can be achieved by disconnecting the left renal vein at the hilum of the kidney and swinging it upward to anastomose end-to-side to the splenic vein (11, 12). If the latter technique is elected, the splenic vein must then be ligated on its mesenteric end to assure compartmentalization of the portal system.

Postoperatively, a nasogastric tube may be utilized to decompress the stomach if preferred. Hourly monitoring of the urinary output and central venous pressure are important to successful management of these patients as they may sequester ascitic fluid in the abdomen for a few days and thus become hypovolemic.

RESULTS

Follow-up experience with 76 personal cases of distal splenorenal shunt suggests that the initial goals of this new procedure are being realized. Late patency of the splenorenal anastomosis has been proven by autopsy study or angiography in more than 90% of postoperative patients in the authors' series. In addition, resolution of varices by roentgenography has equalled the rates achieved with other standard shunts, and prevention of recurrent variceal bleeding has been excellent, with the longest survivors being more than eight years after operation.

Routine postoperative angiography has demonstrated continued portal venous perfusion of the liver in over 90% of patients with patent shunts. Of even greater importance clinically is the fact that careful studies using the maximum rate of urea synthesis [a new index of hepatic functional activity described by Rudman et al (13)] have shown that distal shunt patients retain significantly better metabolic function than randomized controls who had undergone various total shunting procedures (14). Ammonium tolerance and protein tolerance studies are also distinctly superior to values obtained from total shunted patients in our hands. These metabolic data are further supported by the clinical observation that hepatic encephalopathy is a rare complication after the distal shunt procedure (10).

Admittedly, an initial high operative mortality (7 deaths in 12 patients) was observed during the time when the distal shunt operation was being developed and selection criteria being established (3). Over the past several years, however, in experience with 64 selected patients, the operative mortality has been only 5%, a figure which compares very favorably with reports for all other types of shunt procedures. The ultimate assessment of operative result still remains to be determined, namely the important question of long-term survival in respect to other surgical procedures. It is anticipated that prospective, randomized, controlled studies currently under way at Emory will provide conclusive information regarding long-term survival date within a few years (14).

SUMMARY

The selective distal splenorenal shunt has been shown to be equally as effective as total portal-systemic shunting procedures in the control of variceal hemorrhage and it can now be performed with an equivalent operative mortality. It is superior to other shunts in metabolic aspects and results in a markedly decreased incidence of postoperative hepatic encephalopathy. Massive ascites constitutes a definite contraindication to the selective distal splenorenal anastomosis, but for all other suitable patients this procedure appears to be the optimal method of surgical management at this time.

Literature Cited

1. Whipple, A. O. 1946. The rationale of portacaval anastomosis. *Bull. NY Acad. Med.* 22:251–53
2. Blakemore, A. H. 1946. Portacaval anastomosis: A report on 14 cases. *Bull. N.Y. Acad. Med.* 22:254–63
3. Salam, A. A. et al 1971. Hemodynamic contrasts between selective and total portal-systemic shunting. *Ann. Surg.* 173:827–44
4. Callow, A. D. et al 1970. Conclusions from a controlled trial of the prophylactic portacaval shunt. *Surgery* 67:97–103
5. Conn, H. O., Lindenmuth, W. W. 1965. Prophylactic portacaval anastomosis in cirrhotic patients with esophageal varices. *N. Engl. J. Med.* 272:1255–63
6. Jackson, F. C., Perrin, E. B., Felix, W. R., Smith, A. G. 1971. A clinical investigation of the portacaval shunt. V. Survival analysis of the therapeutic operation. *Ann. Surg.* 174:672–701
7. Sherlock, S., Hourigan, K., George, P. 1970. Medical complications of shunt surgery for portal hypertension. *Ann. N.Y. Acad. Sci.* 170:392–405
8. Warren, W. D., Zeppa, R., Fomon, J. J. 1967. Selective trans-splenic decompression of gastroesophageal varices by distal splenorenal shunt. *Ann. Surg.* 166:437–55
9. Price, J. B. Jr., McCullough, W., Peterson, L., Britton, R. C., Voorhees, A. B. Jr. 1967. Effects of portal systemic shunting on intestinal absorption in the dog and in man. *Surg. Gynecol. Obstet.* 125:305–10
10. Warren, W. D., Salam, A. A., Hutson, D., Zeppa, R. 1974. Selective distal splenorenal shunt. Technique and results of operation. *Arch. Surg.* 108:306–14
11. Warren, W. D., Salam, A. A., Faraldo, A., Hutson, D., Smith, R. B. III 1972. End renal vein-to-spenic vein shunts for total or selective portal decompression. *Surgery* 72:995–1006
12. Warren, W. D., Salam, A. A., Smith, R. B., III. 1974. The meso-spleno-renal shunt procedures: A comprehensive approach to portasystemic decompression. *Ann. Surg.* 179:791–98
13. Rudman, D. et al. 1973. Maximal rates of excretion and synthesis of urea in normal and cirrhotic subjects. *J. Clin. Invest.* 52:2241–49
14. Warren, W. D. et al. 1974. The metabolic basis of portasystemic encephalopathy and the effect of selective vs. nonselective shunts. *Ann. Surg.* 180: 573–79

Copyright 1975. All rights reserved

TREATMENT OF TESTICULAR CANCER[1]

❖7146

Donald I. Twito, M.D. and B. J. Kennedy, M.D.
Section of Medical Oncology, Department of Medicine, University of Minnesota School of Medicine, Minneapolis, Minnesota 55455

Carcinoma of the testis is one of the most common forms of lethal malignancy encountered in males 15 to 40 years of age (Table 1). The average annual age-adjusted incidence rate for all races is 3.1 per 100,000 males (1). Although occurring at all ages, it ranks first in incidence rate in males 20 to 34 years of age and accounts for 11.4% of all cancer deaths in the 25–34 age group (2). Approximately 1% of all malignant tumors in males originate in the testis, accounting for 0.6% of all fatal tumors in males (2). The probability at birth of developing cancer of the testis is 0.16% compared to 5.5% for female breast cancer and 2.5% for lung cancer (2). It is nearly three times more common in whites than non-whites.

Cancer of the testis has a significant impact on the social, economic, and emotional status of this young population. There is no doubt that the prevalence of this type of cancer has been underemphasized and the public uninformed. The tragedy in many patients is that approximately six months elapse before definitive diagnosis is made and treatment begun (3–5). It is apparent that education of both patients and physicians is the major means of reducing this interval.

Table 1 Rank order of the most frequent malignancies according to annual incidence rates (per 100,000) in the United States, 1969–71 (1)

	Age (years)				
Rank	15–19	20–24	25–29	30–34	35–39
1	Hodgkins 3.1	TESTIS 6.1	TESTIS 5.8	TESTIS 7.2	Lung 11.8
2	Leukemia 2.8	Hodgkins 5.4	Hodgkins 5.4	Melanoma 4.9	TESTIS 7.0
3	TESTIS 2.0	Leukemia 3.0	Melanoma 3.1	Hodgkins 4.0	Melanoma 5.4

---

[1] This research was supported by Research Grants CA-08101, CA-05158, and CA-08832 from the National Cancer Institute, the Minnesota Medical Foundation, and the Masonic Memorial Hospital Fund Inc.

Once the diagnosis of testis cancer is established, clinical staging is undertaken to determine the extent of the disease in order to select the mode of further therapy and establish a prognosis. It is clear that the extent of spread must be correlated with the histologic tumor type in making comparative analysis of therapeutic modalities. These two factors have been shown to be the most essential in the prognosis of patients with testis cancer (5).

DIAGNOSIS

All investigators agree that patients with suspected testicular cancer should undergo inguinal orchiectomy with ligation of the cord at the internal ring. This serves as local control of the primary tumor and provides adequate material for accurate histologic diagnosis. Multiple serial sections should be made from the specimen, since more than one cell type may be present, influencing prognosis.

Transcrotal biopsy is absolutely contraindicated and needling of the occasionally associated hydrocele should be avoided. Scrotal implantation secondary to these procedures allows the tumor access to the inguinal lymphatics, complicating definitive treatment. Biopsy of distant metastases is recommended since the histology of metastatic lesions may not correspond to that of the primary tumor (6, 7).

PATHOLOGY

Universal agreement is lacking regarding the histologic classification of testicular tumors. The prevailing concept is based on the earlier work of Friedman & Moore (8) and Dixon & Moore (9), recently modified by Mostofi (10). It is presently referred to as the Armed Forces Institute of Pathology (AFIP) Classification, but it is the tentative World Health Organization (WHO) Classification. It separates tumors into two general categories: germinal and nongerminal. Germinal tumors constitute approximately 95% of all testis tumors and are of four basic histologic patterns: seminoma, embryonal carcinoma, teratoma, and choriocarcinoma. A pure cell type is displayed in 60% of testicular neoplasms, and 40% show more than a single cell pattern. The resulting 15 possible combinations of tumor types have been condensed into five clinico-pathologic groups based on two features of clinical behavior (6, 9): The presence of seminoma does not adversely affect the prognosis of other combinations; the presence of teratoma seems to ameliorate the behavior of embryonal carcinoma or choriocarcinoma (Table 2). Nongerminal tumors account for the remaining 5% of testicular neoplasms and include gonadal stromal tumors (3–4%), adnexal tumors, and metastatic tumors. Gonadal stromal tumors are only rarely malignant (11).

Another histological system for tumor classification was proposed by Collins & Pugh in 1964 (12), based on a study of 1115 testicular masses reviewed through the British Tumor Registry. It differs basically from the AFIP system in considering the origin of the tumors. It recognizes only seminomas as of germinal origin. The remainder of testicular tumors of germ cell origin in the AFIP classification are

Table 2 Classification of testicular tumors

Group	Pathologic Type	Distribution Percentage
I	Seminoma, pure	40
II	Embryonal carcinoma, pure or with seminoma	30
III	Teratoma, pure or with seminoma	9
IV	Teratoma, with either embryonal carcinoma or choriocarcinoma or both, and with or without seminoma	20
V	Choriocarcinoma, pure or with either seminoma or embryonal carcinoma or both	1

termed teratoma or malignant teratomas of varying degrees of differentiation. They are considered to be of uncertain histogenesis.

STAGING EVALUATION

To establish the extent of disease, minimum studies include a complete physical examination, chest X ray and whole lung tomograms, biochemistries, intravenous pyelogram, lymphangiography via the pedal and/or inguinal routes, and quantitation of urinary gonadotropins. On the basis of these findings, patients can be placed into three major categories according to the Walter Reed Staging System (Table 3). In some patients final staging is accomplished only after surgical resection of retroperitoneal lymph nodes. Buck et al (13) have advocated left supraclavicular node biopsy in all Stage II patients, since 16% (4 of 25) of their patients had positive biopsies, thereby reclassifying them in the Stage III category.

The most difficult problem with today's methods of staging is the inability to accurately determine the presence or absence of microscopic foci of metastatic cancer in retroperitoneal nodes. To detect these foci, attempts have been made to improve the technique of lymphangiography. The lymphatic drainage of the testis was accurately determined around the turn of the century (14). More recently Chiappa et al (15) described a new technique for visualizing the specific testicular lymph center lateral to but at the level of the first and second lumbar vertebrae on

Table 3 Walter Reed Staging System for testicular neoplasms

Stage IA	Tumor confined to the testis. No clinical or radiographic evidence of spread beyond.
Stage IB	Similar to Stage IA, but found to have histologic spread of tumor at the time or retropertioneal dissection.
Stage II	Clinical or radiographic evidence of metastases to femoral, inguinal, iliac, or para-aortic lymph nodes: no demonstrable metastases above the diaphragm or to visceral organs.
Stage III	Clinical or radiographic evidence of metastases above the diaphragm or other distant metastases to body organs.

the left and the first to third lumbar vertebrae on the right. Since this center was not usually visualized with pedal lymphangiography (16) alone, it was concluded that combined testicular and pedal lymphangiography is the most accurate method of detecting retroperitoneal metastases. Sayegh et al (17), who developed a similar technique, realistically concluded that its most practical contribution had been to the surgeon. They feel that this technique clearly delineates the lower margin of retroperitoneal lymphadenectomy and makes imperative a bilateral node dissection extending to L_1. Since these techniques are still not absolutely descriptive of the extent of microscopic metastases, Stage I patients are managed as Stage II patients and undergo retroperitoneal lymph node dissection.

TREATMENT

The initial attack against this disease is by inguinal orchiectomy with ligation of the cord structures at the internal ring for control of the primary tumor. The histology of the lesion determines the subsequent management of the patient. Orchiectomy was curative in a small percentage of patients (18). Five year survival rates following this operation alone ranged from 5 to 49%. These unsatisfactory results prompted numerous investigators to employ more aggressive modes of therapy for the germ cell neoplasms—primarily irradiation to and resection of the retroperitoneal areas.

Surgery and Irradiation

SEMINOMA The treatment of choice for this germ cell neoplasm has been well established. There is no question that retroperitoneal surgery is unnecessary according to a number of investigators (19–22). Accepted therapy includes inguinal orchiectomy followed by irraditation to the lymph node bearing areas. The only disagreement concerns the extent of irradition necessary. Maier et al (20) and Kurohara et al (4) advocate treating the mediastinum and left supraclavicular area in addition to the para-aortic and ipsilateral iliac lymph nodes. Johnson (19), Castro & Gonzalez (23), and Ytredal & Bradfield (24) believe that treating the subdiaphragmatic areas is sufficient. There is agreement among all these authors that the contralateral iliac area should also be included in the radiation field if the patient has undergone any inguinal surgery prior to development of his malignancy. The lymphatic drainage becomes abnormal following this type of surgery (25–27).

Friedman & Purkayastha (28) have shown that irradiation is also effective in recurrent or disseminated seminoma. They reported occasional cures with radiotherapy even in the presence of massive lymph node metastases or wide spread hematogenous metastates to the lungs and bone. The five year survival rate after recurrence was 55% with 50% survival at 10 years.

Overall survival with seminoma following orchiectomy and irradiation has ranged from 63 to 91% in a number of reports since 1958 (29). Most of these patients have been followed for at least five years. These figures point out that this particular tumor is one of the most curable of all malignancies. Nefzger & Mostofi (30) have shown in a large group of patients with seminoma that those who survive two years after initial therapy without recurrence approach the survival of normal U.S. males.

NONSEMINOMATOUS GERM CELL NEOPLASMS In the treatment of nonseminomatous germ cell neoplasms the roles of retroperitoneal dissection and/or radiotherapy are not clearly defined. Irradiation following inguinal orchiectomy, as the only form of therapy of these tumors, has produced survival ranging from 12 to 53% (2, 29). The wide variation in results is due to the inclusion of patients with both Stage I and Stage II disease and the use of differing end points in reporting survival. In an effort to improve these results, retroperitoneal lymph node dissection has been employed.

The first large series of patients treated by retroperitoneal node dissection attained a five year survival of 46% (31). Since that study, other reports of survival have ranged from 44 to 85%, with recent reports of greater than 65% survival after five years. (29). The true value of retroperitoneal dissection is still difficult to assess, because most of the patients in these series received postoperative irradiation when metastatic disease was found in the retroperitoneal lymph nodes. This pattern of therapy for nonseminomatous germ cell tumors is popular in most major U.S. treatment centers (32). Staubitz et al (33) performed retroperitoneal dissection alone in the treatment of Stage I and Stage II nonseminomatous tumors. Of 53 patients followed for five years or more, there was an 86% survival in Stage I disease and 70% survival in Stage II disease. The transperitoneal approach was employed in this study for performing the retroperitoneal node dissection, emphasizing the importance of bilateral removal of the lymph nodes. This is the only large series where surgery alone has been the mode of therapy. Other authors advocate a unilateral lymph node dissection through a thoraco-abdominal incision (29, 34), claiming that dissection of tissue above the renal pedicle is possible. Survival figures are similar to those for bilateral dissection.

Another prospective study attempting to define the roles of surgery and irradiation is underway at the Walter Reed General Hospital (22, 35). Their patients are randomized, following orchiectomy, between irradiation of the retroperitoneal lymph nodes alone and preoperative irradiation followed by bilateral retroperitoneal lymphadenectomy and postoperative irradiation. It is apparent that further data in such a prospective study for Stage IB and Stage II are also needed for retroperitoneal node dissection alone. Besides survival from these procedures, consideration should be made of the complications of radiotherapy and its long term implications.

For Stage III testis cancer the surgical management of lung metastases has been attempted (36–38). Most patients have had solitary lung nodules although as many as seven nodules were resected in one patient. Survival following this procedure has not been impressive, although a few long term survivors have been noted.

Chemotherapy

Chemotherapy has generally been reserved for those patients with disseminated (Stage III) testicular neoplasms. Histology has been shown to be an important factor in selecting the agent or combination of agents to be employed. Survival of these patients has been increased with some investigators contending that cure has been accomplished.

NONSEMINOMATOUS GERM CELL NEOPLASMS In 1960 Li et al (39) introduced the first major thrust of chemotherapy in advanced testis cancer with the combination of an alkylating agent, an antimetabolite, and an antitumor antibiotic. Of 23 patients, 12 (53%) had objective improvement, 7 with complete disappearance of tumor. From this report the combination of actinomycin D, methotrexate, and chlorambucil evolved and was termed "triple therapy." Subsequent reports confirmed a response rate of 50 to 70%, with 5 to 12% rendered clinically free of disease for varying lengths of time (7, 40, 41).

Mackenzie (7) in 1966 reported on the use of actinomycin D alone in 22 patients with 66% responding, 20% with a complete response. He concluded that actinomycin D was the most effective chemotherapeutic agent for embryonal carcinoma, teratocarcinoma, and choriocarcinoma. It has been implied that the other two agents of the triple therapy may not be as important as considered initially.

Mithramycin was first reported in 1960 to be effective in disseminated testicular tumors (42, 43). These preliminary observations led to the studies of Kofman et al (44) and Brown & Kennedy (45), who concluded that mithramycin was specifically effective in embryonal carcinoma of the testis. Kofman noted responses in 5 (55%) of 9 patients with embryonal cancer of the testis and Brown and Kennedy noted improvement in 7 (58%) of 12 patients. Occasional patients were rendered completely free of disease for prolonged periods of time, the longest 15 years. In spite of these favorable reports, widespread use of this agent was not employed due to hepatic and renal toxicity (46). Kennedy (47) in 1967 detailed an alternate day plan for administration of mithramycin which proved to avoid these toxicities.

Other single agents have been reported in the therapy of nonseminomatous germ cell tumors. Methotrexate produced long term improvement in 4 of 10 patients, all with embryonal cancer (48). Vinblastine produced 2 partial responses in 9 patients (49), whereas Samuels & Howe (50) obtained 11 responses in 21 patients, 4 of which were complete.

Because of the effect of "triple therapy," other combinations of agents have been utilized. Nitrogen mustard and actinomycin D in 14 patients, nitrogen mustard and methotrexate in 15 patients, and vincristine plus an alkylating agent and methotrexate in 12 patients produced response rates of approximately 35% (51).

Later, Jacobs reported that vincristine, actinomycin D, and either cyclophosphamide or cyclophosphamide alternated with mithramycin produced responses in about 50% of patients treated (52). With vincristine, cyclophosphamide, methotrexate, and 5-fluorouracil (COMF), 17 patients had a 41% response rate (53). The use of the combination of vinblastine and melphalan by Samuels & Howe (50) produced a response rate of 45%. Recently the combination of bleomycin and vinblastine or the combination of bleomycin with cyclophosphamide, vincristine, methotrexate, and 5-fluorouracil produced improvement in 75% of 70 patients with a 31% complete response rate (54). The mean survival of the complete responders is at present approximately two years while that of the partial responders is about nine months. Spigel & Coltman (55) reported an 81% response rate in 11 patients with the use of bleomycin and vinblastine. A complete disappearance of tumor occurred in 5 (45%) of these 11 patients. They are presently being maintained on intermittent vinblastine therapy. The response rates of the combinations employing bleomycin are the most impressive to date.

Newer agents continue to be developed for the management of advanced cancer of the testis. The use of adriamycin in 20 patients produced slight improvement in 65%, but only 3 patients had more than a 50% reduction in measurable tumor (56). It was concluded that adriamycin might be more effective when used in a combination drug program.

Most recently, platinum compounds have been used in the treatment of advanced germinal neoplasms. The observation of Rosenberg et al (57) in 1965 that *Escherichia coli* cell division was inhibited by platinum electrode products culminated in the report by Higby et al (58) of platinum's efficacy in human testis cancer. Improvement occurred in 9 of 11 patients, 3 with complete regression of the disease. Regressions have been of rather short duration and renal toxicity is frequent. The study of the combination of platinum and adriamycin is underway.

SEMINOMA Data on the chemotherapy of disseminated seminoma are based on small numbers of patients, since this tumor responds so well to radiotherapy. Mackenzie (7) reported two complete and two partial responses in 4 patients treated with chlorambucil. Golbey (59) concurred on the use of this alkylating agent. Snyder et al (60) obtained favorable responses in 2 patients treated with cyclophosphamide. Blokhin et al (61) treated 23 patients with sarcolysin, obtaining 11 complete and 8 partial tumor regressions. Chebotareva (62) had a similar experience. Melphalan (50) and nitrogen mustard (63) also have been employed and several combinations of chemotherapeutic agents have produced similar results (64).

ADJUVANT CHEMOTHERAPY In view of the fact that approximately 25% of patients with Stage I nonseminomatous tumors eventually develop recurrent disease, the early use of chemotherapy has been advocated by a number of authors (44, 53, 59, 65). Ansfield (41) in 1969 reported the use of triple therapy in 13 patients with nonseminomatous testicular tumors as an adjuvant to orchiectomy, whether or not a retroperitoneal node dissection had been performed. At the time of his report, 11 patients were alive and free of disease, including 3 of 4 patients who had positive nodes at the time of retroperitoneal dissection. Whitmore (6) administered actinomycin D alone or in combination with chlorambucil at one- or two-month intervals for two to three years following retroperitoneal surgery regardless of the histopathology of the resected nodes. Skinner & Leadbetter (29) used actinomycin D as an adjuvant to surgery in patients with positive retroperitoneal nodes. Patients are given the drug during the operative period, the postoperative period, and then at three- to four-month intervals for at least two years. The concept of adjuvant chemotherapy is realistic in view of the complete regressions attained in some patients with disseminated disease.

CONCLUSION

The attack on testicular cancer involves multidisciplinary management including surgery, radiotherapy, and chemotherapy. All play significant roles, but the specific sequence in which they are to be utilized is still evolving. There is no doubt that testicular cancer is one of those neoplasms in which true cure of the disease can be envisioned. A strong program of public education aimed at earlier recognition and detection of the disease would augment the potential results of the therapeutic modalities.

Literature Cited

1. National Cancer Institute, National Institutes of Health. 1974. *Third National Cancer Survey, Advance Three Year Report, 1969–1971 Incidence (Dept. HEW Publ. No. (NIH) 74–637)*
2. MacKay, E, N., Sellers, A. H. 1966. *Can. Med. Assoc. J.* 94:889–99
3. Borski, A. A. 1973, *Cancer* 32:1202–5
4. Kurohara, S. S., George, F. W. III, Dykhuisen, R. F., Leary, K, L. 1967. *Cancer* 20:1089–98
5. Kurohara, S. S., Badib, A. O., Webster, J. H., Martin, L. S. J., Woodruff, M. W. 1968. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 103:827–36.
6. Whitmore, W. F. Jr. 1968. *Cancer Management,* 347–55. Philadelphia: Lippencott. 662 pp.
7. Mackenzie, A. R. 1966. *Cancer* 19: 1369–76
8. Friedman, N. B., Moore, R. A. 1946. *Mil. Surg.* 99:573–93
9. Dixon, F. J., Moore, R. A. 1953. *Cancer* 6:427–54
10. Mostofi, F. K. 1973. *Cancer* 32:1186–1201
11. Mahon, F. B. Jr., Gosset, F., Trinity, R. G., Madsen, P.O. 1973. *Cancer* 31:1208–12
12. Collins, D. M., Pugh, R. C. B. 1964. *Brit. J. Urol. Suppl.* 36:1–11
13. Buck, A. S., Schamber, D. T., Maier, J. G., Lewis, E. L. 1972. *J. Urol.* 107:619–21
14. Jamieson, J. K., Dobson, J. F. 1910. *Lancet* 1:493–95
15. Chiappa, S., Galli, G., Barbaini, S., Ravasi, G. 1962. *Radiol. Med.* 48:663–92
16. Chiappa, S., Uslenghi, C., Bonadonna, G., Marano, P., Ravasi, G. 1966. *Surg. Gynecol. Obstet.* 123:10–14
17. Sayegh, D., Brooks, T., Sacher, E., Busch, F. 1966. *J. Urol.* 95:102–7
18. Vechenski, T. O., Jaeschke, W. H., Vermund, H. 1965. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 95:494–514
19. Johnson, G. C. 1963. *Radiology* 80:539–49
20. Maier, J. G., Mittemeyer, B. T., Sulak, M. H. 1968. *J. Urol.* 99:72–78
21. Dykhuizen, R. F. et al 1968. *J. Urol.* 100:321–28
22. Maier, J. G., Van Buskirk, K. E., Sulak, M. H., Perry, R. H., Schamber, D. T. 1969. *J. Urol.* 101:356–59
23. Castro, J. R., Gonzalez, M. 1971. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 111:355–59
24. Ytredal, D. O., Bradfield, J. S. 1972. *Cancer* 30:628–33
25. Gilbert, J. B. 1941. *J. Urol.* 46:740–97
26. Witus, W. S., Sloss, J. H., Valk, W. I. 1959. *J. Urol.* 81:669–71
27. Bowles, W. T. 1962. *J. Urol.* 88:266–67
28. Friedman, M., Purkayastha, M. C. 1960. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 83:25–42
29. Skinner, D. G., Leadbetter, W. F. 1971. *J. Urol.* 106:84–93
30. Nefzger, M. D., Mostofi, F. K. 1972. *Cancer* 30:1225–32
31. Lewis, L. G. 1953. *J. Urol.* 69:841–44
32. Maier, J. G., Sulak, M. G. 1973. *Cancer* 32:1217–26
33. Staubitz, W. J., Early, K. S., Magoss, I. V., Murphy, G. P. 1973. *Cancer* 32:1206–11
34. Walsh, P. C., Kaufman, J. J., Coulson, W. F., Goodwin, W. E. 1971. *J. Am. Med. Assoc.* 217:309–12
35. Maier, T. G., VanBuskirk, K. E. 1970. *J. Am. Med. Assoc.* 213:97–98
36. Alexander, J., Haight, C. 1947. *Surg. Gynecol. Obstet.* 85:129–46
37. Skinner, D. G., Leadbetter, W. F., Wilkins, E. W. Jr. 1971. *J. Urol.* 105:275–82
38. Thomford, N. R., Woolner, L. B., Clagett, O. T. 1965. *J. Thorac. Cardiov. Surg.* 49:357–63
39. Li, M. C., Whitmore, W. F., Golbey, R., Grabstald, H. 1960. *J. Am. Med. Assoc.* 174:1291–99
40. Moore, C. A. 1968. *J. Urol.* 100:527–29
41. Ansfield, F. J., Korbitz, B. C., Davis, H. L. Jr. 1969. *Cancer* 24:442–46
42. Curreri, A. R., Ansfield, F. J. 1960. *Cancer Chemother. Rep.* 8:18–22
43. Parker, G. W., Wiltsie, D. S., Jackson, C. B. Jr. 1960. *Cancer Chemother. Rep.* 8:23–26
44. Kofman, S., Medrek, T. J., Alexander, R. W. 1964. *Cancer* 17:938–48
45. Brown, J. H., Kennedy, B. J. 1965. *N. Engl. J. Med.* 272:111–18

46. Kennedy, B. J. 1970. *Am. J. Med.* 49:494–503
47. Kennedy, B. J. 1970. *Cancer* 26:755–66
48. Wyatt, J. K., McAninch, L. N. 1967. *Can. Med. Assoc. J.* 10:421–26
49. Warwick, D. H., Alison, R. E., Darte, J. M. N. 1961. *Can. Med. Assoc. J.* 85:579–83
50. Samuels, M. L., Howe, C. D. 1970. *Cancer* 25:1009–19
51. Jacobs, E. M., Johnson, F. D., Wood, D. A. 1966. *Cancer* 19:1697–1704
52. Jacobs, E. M. 1970. *Cancer* 25:324–26
53. Medelsohn, D., Serpick, A. A. 1970. *J. Urol.* 103:619–23
54. Samuels, M. L. 1974. *Bleomycin in the therapy of testicular tumors.* Presented at the New Drug Seminar on Bleomycin sponsored by the National Cancer Institute, National Institutes of Health, Bethesda
55. Spigel, S. C., Coltman, C. A. Jr. 1974. *Cancer Chemother. Rep.* 58:213–16
56. Monfardini, S., Bajetta, E., Musumeci, R., Bonadonna, G. 1972. *J. Urol.* 108:293–96
57. Rosenberg, B., Van Camp, L., Krigas, T. 1965. *Nature* 205:698–99
58. Higby, D. J., Wallace, H. J. Jr., Holland, J. F. 1973. *Cancer Chemother. Rep.* 57:459–63
59. Golbey, R. B. 1970. *J. Am. Med. Assoc.* 213:101–3
60. Snyder, W., Rodensky, P., Lieberman, B. 1964. *Cancer Chemother. Rep.* 41:37–40
61. Blokhin, N., Larionov, L., Perevodchikova, N., Chebotareva, L., Merkulova, N. 1958. *Ann. N.Y. Acad. Sci.* 68:1128–32
62. Chebotareva, L. 1960. *Abstr. Vop. Onkol.* 6:108
63. Cline, M. J. 1971. *Cancer Chemotherapy,* 54–57. Philadelphia: Saunders. 205 pp.
64. duPriest, R. W. Jr., Fletcher, W. S. 1973. *Oncology* 28:147–63
65. Ream, N. W., Perlia, C. P., Wolter, J., Taylor, S. G. III. 1968. *J. Am. Med. Assoc.* 204:1030–36

Copyright 1975. All rights reserved

MUSCLE AMINO ACID METABOLISM AND GLUCONEOGENESIS

❖7147

Neil B. Ruderman, M.D.[1]
Joslin Research Laboratory, Department of Medicine, Harvard Medical School, and the Peter Bent Brigham Hospital, Boston, Massachusetts 02115

INTRODUCTION

In addition to its role in locomotion, muscle, by virtue of its mass, is the body's major reservoir of amino acids. During a fast, amino acids derived from muscle protein are an important substrate for gluconeogenesis (1, 2); in fact, recent investigations suggest that during prolonged starvation (3), and in certain hypoglycemic disorders (4), hepatic glucose production is depressed because the supply of amino acids from muscle is decreased. It is the aim of this review to explore the relationship between amino acid metabolism in muscle and hepatic gluconeogenesis in fasting man. In addition I will comment briefly on the role of substrate supply in disorders of gluconeogenesis and will describe some therapeutic approaches which have recently been used to limit amino acid catabolism in burn victims and in patients with renal failure. For a more comprehensive discussion of selected aspects of this rapidly burgeoning field, the reader is referred to the series *Mammalian Protein Metabolism* (5–11), and to the reviews on gluconeogenesis by Exton (12), and alanine metabolism by Felig (13).

AMINO ACID RELEASE FROM SKELETAL MUSCLE

Alanine

Although it has long been realized that amino acids are released by muscle, it has only recently been appreciated that the mixture of amino acids released does not reflect the composition of muscle protein. London et al (14) and later Pozefsky and his co-workers (15) noted that alanine accounts for a far greater percentage of the amino acid released by the human forearm than would be expected from its content

---

[1] Supported in part by National Institutes of Health Grant Am-15272 and by National Institutes of Health Research Career Development Award Am-15729.

Table 1 Relation between the pattern of amino acid release from the human forearm and the amino acid composition of muscle

Amino Acid	Percentage of Amino Acid Released by Human Forearm[a]	Percentage of Muscle Protein[b]
Alanine	28	6.4
Glutamine	23	6.6
Lysine	8.7	12.6
Glycine	7.9	4
Proline	6.7	2.4
Arginine	4.4	13.5
Threonine	4.4	3.9
Phenylalanine	2.8	2.3
Histidine	2.8	3.1
Valine	2.8	3.5
Tyrosine	2.4	1.9
Leucine	2	6.2
Isoleucine	2	3.9
Aspartate	0	7
Glutamate	-7.9 (net uptake)	11.7

[a] Data of Aoki et al (16), based on arterio-deep venous differences in whole blood in the postabsorptive state.

[b] A theoretical protein containing actin and myosin in the proportion they exist in skeletal muscle. Adapted from Kominz et al (17).

in muscle protein (Table 1). Since muscle does not contain a labile protein high in alanine, but possesses glutamic pyruvic transaminase in high activity (10), it was reasoned that the additional alanine is probably formed by transamination of pyruvate. In accordance with this view, it has been shown that infusing pyruvate into the brachial artery of man causes an increase in alanine efflux (18). Likewise, alanine formation has been demonstrated in rat muscle by perfusing it either with lactate, which increases tissue pyruvate, or with leucine, which transfers its amino group to α-ketoglutarate and causes a marked increase in muscle glutamate (19, 20; see Figure 1).

The Cori Cycle and the Glucose-Alanine Cycle

The origin of the pyruvate from which alanine is formed has not been completely resolved. Cori pointed out that some of the glucose and glycogen utilized by peripheral tissues is not oxidized, but is released as lactate and pyruvate, which are reconverted to glucose by the liver and then cycled back to the periphery (21). The tissues which contribute to the Cori cycle include red cells, skin, and renal medulla, which derive energy principally from anaerobic glycolysis, and skeletal muscle and brain which, during a fast, release increased quantities of lactate and pyruvate due to inhibition of pyruvate oxidation. Felig (13) and Mallette, Exton & Park (22) have

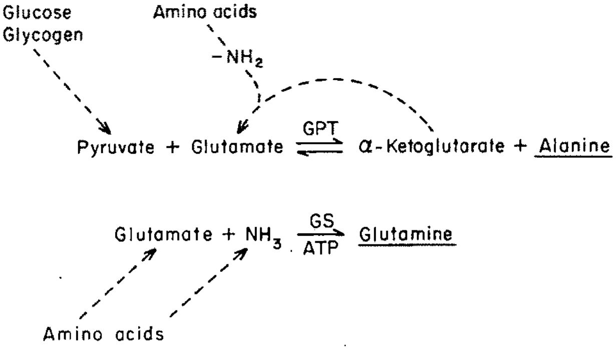

Glucose
Glycogen
Amino acids
$-NH_2$
GPT
Pyruvate + Glutamate ⇌ α-Ketoglutarate + Alanine
Glutamate + NH_3 $\xrightarrow[ATP]{GS}$ Glutamine
Amino acids

Figure 1 Alanine and glutamine formation in skeletal muscle. *GPT* = glutamic pyruvic transaminase; *GS* = glutamine synthetase. Alanine formation is favored by an increase in muscle pyruvate or glutamate and glutamine release by an increase in NH_3 or glutamate (19, 20).

suggested that the alanine released from muscle is part of a similar cycle, a "glucose-alanine cycle," in which the carbon required for alanine formation is derived principally from exogenous glucose and muscle glycogen, and the nitrogen from the catabolism of other amino acids.

Insofar as it is derived from exogenous glucose, alanine does not contribute to net glucose production. On the other hand, during the first 4–5 days of a fast, a considerable proportion of the lactate, pyruvate, and alanine generated by muscle could be derived from muscle glycogen.[2] The data of Hultman & Bergstrom (23) suggest that as much as 140 g of muscle glycogen, nearly twice the amount provided by liver, are utilized during the first 5 days of a total fast (Table 2). As shown in Figure 2, this could be of considerable importance since gluconeogenesis has to be at a maximum during this period, as hepatic glycogen is expended by day 2 and ketone bodies probably do not become a significant fuel for brain until at least the 4th–5th day, when their concentrations are substantially increased (27; see Table 3). Parenthetically, it is during the first days of a fast that the release of alanine, lactate, and pyruvate from muscle is greatest (3, 29; see Table 3).

It has been suggested that the pyruvate from which alanine is formed may also be derived from amino acids. Theoretically, muscle could generate pyruvate from glutamate, aspartate, and other amino acids which are metabolized in the citric acid cycle by way of a reaction catalyzed by phosphoenolpyruvate carboxykinase (33); however, the quantitative importance of this pathway remains to be assessed.

Glutamine

Like alanine, glutamine is released from muscle in quantities beyond that predicted from its content in muscle protein (24; see Table 1). Skeletal muscle possesses

---

[2] Muscle does not possess glucose-6-phosphatase; therefore, unlike liver and kidney, it cannot convert glucose-6-phosphate derived from glycogenolysis into free glucose.

Table 2 Gluconeogenic substrate in fasting man[a]

Source	Quantity Available After Overnight Fast (g)	Quantity Expended[c] Days 2–4 (g/day)	Days 24–48
Protein (mainly muscle)	3600[b]	48	16
Muscle Glycogen	420	28	0
Adipose Tissue			
Glyceride-Glycerol	900	17	17

[a] Data are approximate and are extrapolated from studies of Cahill (1,2,25) and Hutman (23,24) and co-workers. Subject is assumed to have 28 kg of muscle and 10 kg of fat. Day 2 begins approximately 36 hr after the last meal.

[b] The size of the protein reservoir needed to supply this much gluconeogenic substrate would be 6300 g, since only 57% of the carbon in muscle protein can be converted to glucose (26).

[c] In addition to gluconeogenesis from amino acids, glycerol, and substrates derived from muscle glycogen, approximately 40 g/day would be derived from exogenous glucose via the Cori cycle. The 90 g of liver glycogen present after an overnight fast is expended by 36 hr of starvation (24).

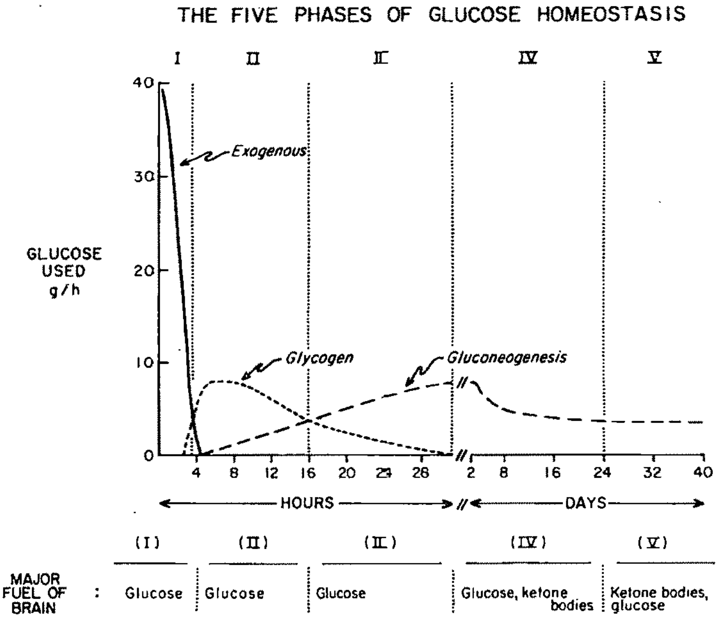


Figure 2 The five phases of glucose homeostasis based on the origin of glucose in blood. A theoretical 70-kg man ingests 100 g of glucose and then proceeds to fast for 40 days. *Phase I* = absorptive; *Phase II* = postabsorptive; *Phase III* = early starvation; *Phase IV* = intermediate starvation; *Phase V* = prolonged starvation. In *Phase V* ketone bodies account for 58% and glucose 28% of the oxidative metabolism of brain. Adapted from Ruderman et al (28).

Table 3 Metabolic parameters in fasting man[a]

	Duration of Fast		
	Overnight	3 Days	28–42 Days
Insulin, μU/ml	37	20	17
Glucagon, pg/ml	75	140	110
Glucose, mM	4.8	3.8	3.8
Acetoacetate + β-hydroxybutyrate, mM	0.1	1.8	7.4
Alanine, μM	344	321	140
Leucine, μM	112	152	72
A–V Difference for alanine, μM			
Forearm	110		30
Splanchnic bed	115		60
Urine-N, g/day	14	11	5

[a] Data were obtained in obese subjects fasted for 5–6 weeks by Cahill and co-workers (see references 3,25,30–32). Amino acid values are for plasma.

glutamine synthetase, an enzyme capable of forming glutamine from glutamate and NH_3 (35; Figure 1), and both our laboratory (20) and that of Hills (36) have demonstrated glutamine formation when muscle of experimental animals is exposed to a high concentration of NH_3. Likewise, we have observed an increase in glutamine formation when rat muscle is perfused with leucine and other amino acids that cause an increase in tissue glutamate when they are deaminated (Figure 1). Furthermore, the increase in glutamine release could be partially suppressed by L-methionine DL-sulfoxamine, an inhibitor of glutamine synthetase (19, 20). Direct evidence that glutamine is formed via glutamine synthetase in human muscle has not yet been obtained; however, in a group of patients with hyperammonemia due to hepatic dysfunction, we have noted a large uptake of NH_3 and an arteriovenous difference across the forearm for glutamine nearly threefold greater than that for alanine (O. P. Ganda and N. B. Ruderman, unpublished data).

In contrast to alanine, the carbon released as glutamine is undoubtedly derived from other amino acids (19), most notably glutamate and aspartate. As shown in Table 1, glutamate and aspartate comprise 18% of muscle protein, but are not released by muscle; in fact, there is invariably a net uptake of glutamate from the circulation. Thus glutamine, in contrast to alanine, is a major vehicle for transporting carbon derived from the breakdown of muscle protein. As will be noted below, only a portion of this glutamine is directly used for gluconeogenesis.

Glutamine and Alanine Metabolism in the Gut

Marliss et al (34) observed that a large amount of glutamine is extracted by the splanchnic bed (liver and gut) and suggested that it ranked with alanine as a major gluconeogenic substrate for liver. This interpretation seemed quite reasonable since experiments with the perfused rat liver had previously revealed that glutamine is a relatively good precursor for hepatic glucose production (37). On the other hand, investigations originally carried out in the dog (38) and since confirmed in man (39) and the rat (40) indicate that most of the glutamine extracted in the splanchnic bed is taken up by gut rather than liver. Furthermore, the gut releases large quantities of alanine—perhaps as much as 50–100% of that released by muscle. Thus earlier assessments of gluconeogenic substrate based on measurements of splanchnic extraction underestimated the importance of alanine and overestimated that of glutamine as hepatic gluconeogenic substrates. Alanine probably accounts for about 50% of the amino acid extracted by the liver in fasted man (see references 39, 41). To what extent this alanine is derived from protein via formation from glutamine in the gut is not known.

Physiological Significance of Alanine and Glutamine Formation

Muscle does not possess the enzymatic apparatus to synthesize urea; therefore, by releasing nitrogen as alanine and glutamine, it is able to utilize amino acids as a fuel without risk of either releasing large quantities of NH_3 into the circulation (12, 13) or trapping an excessive amount of α-ketoglutarate and oxaloacetate as glutamate and aspartate, thereby inhibiting the citric acid cycle. It is not clear whether any advantage is gained by releasing gluconeogenic precursor as alanine rather than as lactate or pyruvate as all three are excellent gluconeogenic substrates (37). Alanine has been shown to stimulate the secretion of glucagon which, in addition to its well known glycogenolytic effect, is a potent stimulus of gluconeogenesis (42). Whether the moderate increase in circulating glucagon which occurs during early starvation (Table 3) is related to the release of gluconeogenic precursor from muscle as alanine rather than as lactate or pyruvate remains to be determined, however.

The release of alanine is more closely related to nutritional state than is that of glutamine. Studies from our laboratory suggest that glutamine release by muscle is comparable in recently fed (3 hr postprandial) and fasted rats, whereas alanine release is negligible in recently fed animals, but is markedly increased after 24–48 hr of starvation (19; M. MacDonald, N. Ruderman et al, unpublished data). This suggests that alanine transports most of the extra nitrogen generated when the catabolism of amino acids in muscle is increased, as it is during a fast. The precise reason alanine fills this role is not clear; conceivably pyruvate is more available for transamination during starvation because pyruvate oxidation is inhibited (43).

The observation that glutamine release is relatively independent of nutritional state may reflect the fact that its sole role is not gluconeogenic. Muscle is the principal supplier of glutamine for the kidney, which uses it to generate NH_3 in order to conserve cation while excreting an acid urine (see reference 36). Although during starvation glutamine is the major gluconeogenic substrate of the kidney, in the rat, at least, it is also extracted in the fed state, when the kidney is presumably not producing glucose (44). Likewise, glutamine is extracted by the gut in both the fed and fasted states, albeit its role here is not known (M. MacDonald, N. Ruderman et al, unpublished data).

Release of Other Amino Acids

In comparison with alanine, the other amino acids released by muscle play a relatively small role in hepatic gluconeogenesis. Most of the glutamine released is taken up by the gut and kidney. Of the others, only glycine, proline, threonine, and serine are extracted by the splanchnic bed in large amounts, and even then the arterio-hepatic-venous difference of the four combined is less than that of alanine (41). Furthermore, most of the serine which enters the circulation originates in kidney rather than muscle (45).

Metabolism of Branched-Chain Amino Acids

Unlike most other essential amino acids, leucine, isoleucine, and valine, the branched-chain amino acids, are oxidized predominantly in muscle rather than liver (10). In rat muscle their oxidation is increased in starvation (46) and other protein catabolic states (47). Odessey (48) has demonstrated that this is probably related to an increased affinity of the deaminated forms of these amino acids (i.e. ketoleucine, ketovaline, and ketoisoleucine) for the ketoacid decarboxylase(s). Thus at a given concentration of a branched-chain amino acid within the muscle cell, its rate of catabolism is greater in the fasted than the fed rat, and more nitrogen is available for alanine formation. Although one would expect this to lower the concentrations of the branched-chain amino acids in muscle, their release is increased in the presence of insulin deficiency (19, 49). Furthermore, high levels of leucine, isoleucine, and valine in plasma are a hallmark of clinical states associated with increased protein catabolism, even when the concentrations of gluconeogenic amino acids, which are predominantly metabolized in liver, are low (see below).[3]

PROTEIN CONSERVATION AND THE ADAPTATION TO STARVATION

During a fast, the release of alanine and other amino acids from muscle is maximal during the first few days when the brain oxidizes predominantly glucose and the demand for gluconeogenesis is greatest (Table 3, Figure 2). With prolonged starvation, ketone bodies progressively replace glucose as a fuel for brain (25), and pari passu, as the need for gluconeogenesis diminishes, the release of amino acids from muscle is decreased (1–3). The net result of these alterations of brain and muscle metabolism is a conservation of muscle protein at the expense of adipose tissue triglyceride which provides the fatty acids for ketogenesis (1, 2). The basis for the decrease in muscle protein degradation is not known (see below); however, teleologically the advantage of this adaptation is clear. Unlike carbohydrate which can be stored as glycogen in liver and muscle, and lipid which is stored as triglyceride in adipose tissue, there is no storage form of protein: every protein molecule appears to have a purpose, whether it be an enzyme, a structural component of tissue such as collagen, or an element of a contractile system, such as actin and myosin in skeletal muscle. Thus a protracted use of body protein for gluconeogenesis results in a loss of functioning tissue and an organism less able to cope with its environment. In this regard nutritional data suggest that if one third to one half of muscle protein is depleted, terminal bronchopneumonia ensues, presumably due to inadequate clearing of the tracheobronchial tree, as a result of protein loss from the diaphragm and intercostal musculature.

---

[3] In part this is due to increased protein degradation in the liver, which releases branched-chain amino acids into the circulation in diabetes (50).

MODULATORS OF AMINO ACID RELEASE

Early Starvation

The magnitude of amino acid release from muscle is primarily determined by the balance between protein synthesis and degradation. As shown in Table 4, many factors including insulin and glucocorticoids, nutritional state, exercise, denervation, and possibly β-adrenergic agents and inhibitors can affect this balance. During a brief fast, the principal modulator appears to be insulin (51). As the glucose concentration diminishes in the first days of starvation (Table 3), the insulin concentration in plasma falls and the release of alanine and other amino acids from muscle increases. Plasma alanine does not increase, since the decrease in insulin, in combination with a moderate or relative increase in glucagon, markedly enhances alanine extraction and conversion to glucose by liver.

Glucocorticoids appear to play only a permissive role in the adaptation of normal man to a fast: they are required for the increase in amino acid release but they do not modulate it. The production of glucocorticoids is not increased during starvation (27). In addition, they only stimulate muscle protein catabolism in man when administered in pharmacological doses or when given to adrenalectomized or insulin-deficient subjects (56, 57). Glucagon which actively stimulates net protein degra-

Table 4 Factors affecting amino acid release and protein turnover in skeletal muscle[a]

Influence	Net Amino Acid Release	Protein Synthesis	Protein Degradation	References
Insulin	Decreased	Increased	Decreased	11,49,51,52
Glucocorticoid	Increased	Decreased	Increased	11,52
Growth hormone		Increased	No change	52
Brief starvation	Increased	Decreased	Increased	53
Prolonged starvation	Decreased			2,3
Disuse or denervation		Decreased	Increased	52
Work-induced hypertrophy		Increased	Decreased	52
β-Adrenergic agents	Decreased			54,55

[a] Blank spaces indicate that the effect of the agent is not known; "No change," that it has no effect.

dation and gluconeogenesis from amino acids in liver (12, 22) probably has no direct effect on amino acid metabolism in skeletal muscle (19). Whether it secondarily enhances amino acid release from muscle by enhancing amino acid extraction by the liver and thus lowering plasma levels is not known.

Prolonged Starvation

The diminution in amino acid release from muscle in prolonged starvation is associated with a decreased rate of proteolysis (58); however, the factor(s) modulating this change are not known. The concentration of insulin is relatively constant (see Table 3) and the other factors listed in Table 4 are not known to change. The decrease in amino acid release occurs at the same time muscle is switching from ketone bodies to free fatty acids as its major fuel (29); however, the relationship between these events is not clear. It is not known whether the capacity of muscle to oxidize branched-chain amino acids is diminished in prolonged starvation. Conceivably this might exert a regulatory influence on protein synthesis by affecting the composition of the amino acid mix within the cell (59).

AMINO ACID METABOLISM IN HEART

The pattern of amino acid release from heart is similar to that of skeletal muscle (49, 60); however, in vivo, the two tissues respond very differently to insulin deficiency. Whereas untreated diabetes is characterized by wasting of skeletal muscle, it is not associated with cardiac atrophy or acute congestive heart failure. Morgan, Jefferson and co-workers (49) have attributed this to the fact that fatty acids and ketone bodies are able to maintain the protein synthetic apparatus in cardiac muscle when insulin is lacking. They noted that skeletal muscle of a diabetic rat contains increased numbers of ribosomal subunits and decreased polysomes and has a diminished rate of protein synthesis, whereas in heart the numbers of subunits and polysomes are the same as in controls, and protein synthesis is not depressed. They also noted that ribosomal disaggregation and the diminished capacity for protein synthesis, which occur when hearts are perfused for several hours with an insulin-free medium, can be prevented by adding fatty acids, ketone bodies, or insulin to the perfusate. In skeletal muscle only insulin was effective. Thus when plasma insulin is low or biologically ineffective and the degradation of skeletal muscle protein is increased to supply amino acids for gluconeogenesis, heart muscle is spared.

SUBSTRATE SUPPLY AND THE REGULATION OF GLUCONEOGENESIS

In a physiological setting, the magnitude of hepatic gluconeogenesis depends upon the rate at which substrate is provided from peripheral tissues as well as the gluconeogenic capacity of the liver. In man this is most evident during prolonged starvation when the ability of the liver to synthesize glucose is not impaired, but the release of alanine and other amino acids from muscle is decreased, their plasma levels fall, and gluconeogenesis is depressed (3). In this regard, exogenous alanine causes a brisk increase in blood glucose when administered to patients fasted 5–6 weeks (61). The importance of substrate supply has also been demonstrated in the perfused rat liver where Exton et al (62) have demonstrated that glucose production can be controlled by variations in the concentrations of lactate, pyruvate, glycerol, and amino acids within the physiological range.

SUBSTRATE SUPPLY AND DISORDERS OF GLUCONEOGENESIS

Several hypoglycemic disorders in man have been attributed to an insufficient provision of gluconeogenic substrate (reviewed by 4, 13, 28) (Table 5). In great measure, this has been due to the finding of decreased concentrations of alanine and other gluconeogenic amino acids in the circulation. In other words, these patients have been likened to individuals undergoing a prolonged fast in whom the concentrations of most gluconeogenic acids in plasma are decreased due to a diminished release from muscle. It must be emphasized, however, that a low concentration of alanine in plasma or blood does not, by itself, indicate that alanine release from peripheral stores is diminished. Very low concentrations of alanine have been observed in patients in diabetic ketoacidosis (73) and in rats made hypoglycemic with phloridzin (67), an agent which blocks the tubular reabsorption of glucose. In both situations, alanine release from muscle is increased and hypoalaninemia results from enhanced alanine extraction by the liver. In contrast to prolonged starvation, these states are

Table 5 Disorders of glucose homeostasis associated with hypoalaninemia[a]

Disorder	Plasma Leucine	Urine-N Excretion	References
Hypoglycemia			
Ketotic hypoglycemia	Increased		63
Pregnancy, starvation	Increased[b]	Increased	64–66
Phloridzin	Increased	Increased	67
Adrenal insufficiency	Decreased		4
Pituitary insufficiency	Decreased		4
Marasmus	Decreased		68–70
Extrapancreatic tumors			71
Hyperglycemia			
Diabetes	Increased		72
Diabetic ketoacidosis	Increased	Increased	73
Trauma	Increased	Increased	

[a] Blanks indicate that information is not known

[b] The increase in plasma leucine is less than in a nonpregnant woman starved for the same period of time. This may be due to the fact leucine is being extracted by the fetus for protein systhesis.

characterized by a high rate of nitrogen excretion in the urine and by another hallmark of increased protein catabolism, a high plasma concentration of the branched-chain amino acids. Presumably the concentrations of the latter are elevated because they are not utilized for gluconeogenesis by the liver and their oxidation in extrahepatic tissues (see above) is not sufficiently increased to compensate for the extra amino acid generated by the degradation of muscle and liver protein.

Ketotic hypoglycemia, the most common hypoglycemic disorder of childhood, has been classified as a disorder of gluconeogenesis due to lack of substrate, both because plasma alanine levels in affected children are low and because alanine administration causes a marked increase in blood glucose (63). On the other hand, these children, when fasted or placed on a low-carbohydrate, hypocaloric diet, have high levels of branched-chain amino acids (63, 74), suggesting that protein catabolism and the provision of gluconeogenic substrate is accelerated rather than depressed. It is noteworthy that a similar set of biochemical alterations occurs in pregnant women during a fast and in rats treated with phloridzin (Table 5). In both situations protein catabolism is increased and substrate supply for gluconeogenesis is inadequate only because of a markedly increased requirement for glucose. In the pregnant woman this is due to the presence of the fetus, and in the phloridzinized rat to glucosuria. It is also noteworthy that alanine administration causes a brisk increase in blood glucose in both of these situations.

THERAPEUTIC APPROACHES TO LIMITING PROTEIN CATABOLISM

Trauma

Following severe trauma, the body's attempts to spare protein are abandoned and protein catabolism in muscle and other tissues is markedly accelerated. As shown by Long et al (75), this is associated with a substantial increase in glucose production and utilization. This may reflect the needs of the reparative tissues which are obligate glycolyzers.

The role of hormonal factors in the response to trauma has been extensively studied (76). Glucocorticoids are required for the protein catabolic response; however, as it can occur without an increase in glucocorticoid production (e.g. in an adrenalectomized patient or experimental animal on maintenance corticoid), their role is probably permissive (76). Allison et al (77) have recently noted that patients with severe burns have a decreased insulin response to intravenous glucose in the acute phase (attributed to high levels of catecholamines), and that this is followed by an extended period of insulin resistance. When they administered high doses of insulin (200–600 U/day), in combination with KCl and sufficient intravenous glucose to maintain normoglycemia to these patients, a protein-sparing effect was observed, potassium loss in the urine was diminished, and many of the patients could be maintained in positive nitrogen balance (78). As noted by the the authors, therapy to suppress the protein catabolic state may be of little benefit in patients with relatively minor trauma; however, in certain instances, e.g. severe burns and trauma, major surgery, or acute renal failure, when negative nitrogen balance and electrolyte changes could cause death before the underlying disorder is corrected, its use could be lifesaving.

Chronic Renal Failure

In patients with severe renal failure the capacity of the kidney to excrete nitrogenous wastes is exceeded even when protein catabolism is not increased. For this reason, such patients have been treated with high carbohydrate diets, containing small amounts of protein (10–15 g) high in essential amino acid content, to limit protein catabolism (79, 80). Recently it has been suggested that nitrogen catabolism might be diminished even further by administering the essential amino acids as their deaminated ketoanalogs (81, 82, reviewed by 83). Since many of the ketoanalogs can be converted to their corresponding amino acids in man (81–84) and experimental animals (85), it was reasoned that they might diminish the nitrogen load for the kidney by using some of the excess NH_3 generated in the uremic state. Furthermore, it has been established that several of the ketoanalogs can be aminated by muscle and liver without interfering with the metabolism of these tissues (85).

The therapeutic use of ketoanalogs in man has been limited. In a preliminary study, Walser et al (82) administered the ketoanalogs of leucine, isoleucine, valine, methionine, and phenylalanine to a group of uremic patients for 15-18 days, and noted improvements in blood urea nitrogen in all ten subjects. In light of these findings and the apparent safety of these compounds, an evaluation of their therapeutic efficacy, particularly as a means of eliminating or diminishing the need for dialysis, is clearly warranted. Theoretically, the ketoanalogs might also be beneficial in patients with liver disease and a propensity to hepatic encephalopathy. To my knowledge, they have not yet been evaluated for this purpose.

Literature Cited

1. Cahill, G. F. Jr., Owen, O. E. 1968. *Carbohydrate Metabolism and its Disorders*, ed. F. Dickens, P. J. Randle, W. J. Whelan, 497–522. New York: Academic
2. Cahill, G. F. Jr. 1970. *N. Engl. J. Med.* 282:668–75
3. Felig, P., Marliss, E., Pozefsky, T., Cahill, G. F. Jr. 1970. *Am. J. Clin. Nutr.* 23:986–92
4. Pagliara, A. S., Karl, I. E., Haymond, M., Kipnis, D. M. 1973. *J. Pediat.* 82:365–79, 558–77
5. Munro, H. N., Allison, J. B., Eds. 1964. *Mammalian Protein Metabolism*, Vol. 1. *Biochemical Aspects of Protein Metabolism.* New York: Academic. 566 pp.
6. Munro, H. N., Allison, J. B., Eds. 1964. *Mammalian Protein Metabolism*, Vol. 2. *Nutritional and Pathological Aspects of Protein Metabolism.* New York: Academic. 642 pp.
7. Munro, H. N., Ed. 1969. *Mammalian Protein Metabolism*, Vol. 3. *Protein Metabolism During Evolution and Development of Mammals and Technical Aspects of the Study of Protein Metabolism.* New York: Academic. 571 pp.
8. Munro, H. N., Ed. 1969. *Mammalian Protein Metabolism*, Vol. 4. *Regulation Mechanisms in Protein Metabolism.* New York: Academic. 763 pp.
9. Munro, H. N. 1969. *Mammalian Protein Metabolism*, ed. H. N. Munro, 4:299–386. New York: Academic
10. Young, V. 1969. *Mammalian Protein Metabolism*, ed. H. N. Munro, 4:585–674. New York: Academic
11. Manchester, K. L. 1969. *Mammalian Protein Metabolism*, ed. H. N. Munro, 4:229–98. New York: Academic

12. Exton, J. H. 1972. *Metabolism* 21:945–90
13. Felig, P. 1973. *Metabolism* 22:179–207
14. London, D. R., Foley, T. H., Webb, C. G. 1965. *Nature* 208:588–89
15. Pozefsky, T., Felig, P., Tobin, J. D., Soeldner, J. S., Cahill, G. F. Jr. 1969. *J. Clin. Invest.* 48:2273–82
16. Aoki, T. T., Brennan, M. F., Muller, W. A., Cahill, G. F. Jr. 1974. *Advan. Enzyme Regul.* 12:3–10
17. Kominz, D. R., Hough, A., Symonds, P., Loki, D. 1954. *Arch. Biochem. Biophys.* 50:148–59
18. Pozefsky, T., Tancredi, R. G. 1972. *J. Clin. Invest.* 51:2359–69
19. Ruderman, N. B., Berger, M. 1974. *J. Biol. Chem.* 249:5500–6
20. Ruderman, N. B., Lund, P. 1972. *Isr. J. Med. Sci.* 8:295–302
21. Cori, C. F. 1931. *Physiol. Rev.* 11:143–275
22. Mallette, L. E., Exton, J. H., Park, C. R. 1969. *J. Biol. Chem.* 244:5724–28
23. Hultman, E., Bergstrom, J. 1967. *Acta Med. Scand.* 182:109–17
24. Nilsson, L. H., Hultman, E. 1973. *Scand. J. Clin. Lab. Invest.* 32:317–23, 325–30
25. Owen, O. E., Felig, P., Morgan, A. P., Wahren, J., Cahill, G. F. Jr. 1969. *J. Clin. Invest.* 48:574–83
26. Krebs, H. A. 1964. *Mammalian Protein Metabolism* ed. H. N.Munro, J. B. Allison, 1:125–76. New York: Academic
27. Cahill, G. F. Jr. et al 1966. *J. Clin. Invest.* 45:1751–69
28. Ruderman, N. B., Aoki, T. T., Cahill, G. F. Jr. 1975. *Gluconeogenesis,* ed. M. Mehlman, R. Hanson, New York: Academic. In press
29. Owen, O. E., Reichard, G. A. 1971. *J. Clin. Invest.* 50:1536–45
30. Owen, O. E. et al 1967. *J. Clin. Invest.* 48:574–83
31. Marliss, E. B., Aoki, T. T., Unger, R. H., Soeldner, J. S., Cahill, G. F. Jr. 1970. *J. Clin. Invest.* 49:2256–70
32. Felig, P., Owen, O. E., Wahren, J., Cahill, G. F. Jr. 1969. *J. Clin. Invest.* 48:584–94
33. Opie, L. H., Newsholme, E. A. 1967. *Biochem. J.* 103:391–99
34. Marliss, E. B., Aoki, T. T., Pozefsky, T., Most, A. S., Cahill, G. F. Jr. 1971. *J. Clin. Invest.* 50:814–17
35. Iqbal, K., Ottoway, J. H. 1970. *Biochem. J.* 119:145–56
36. Hills, A. G., Reid, E. L., Kerr, W. D. 1972. *Am. J. Physiol.* 223:1470–76
37. Ross, B. D., Hems, R., Krebs, H. A. 1967. *Biochem. J.* 102:942–51
38. Elwyn, D. H., Parikh, C., Shoemaker, W. C. 1968. *Am. J. Physiol.* 215:1260–75
39. Felig, P. et al 1973. *Diabetes* 22:573–76
40. Ishikawa, E., Aikawa, T., Matsutaka, H. 1972. *J. Biochem.* 71:1097–99
41. Felig, P., Wahren, J., Rof, L. 1973. *Proc. Nat. Acad. Sci.* 70:1775–79
42. Marliss, E. B., Aoki, T. T., Cahill, G. F. Jr. 1972. *Glucagon: Molecular Physiology: Clinical and Therapeutic Implications,* ed. P. J. Lefebvre, R. H. Unger, 123–50. New York: Pergamon
43. Berger, M., Hagg, S., Ruderman, N. B. 1973. *Diabetes* 22:292 (Abstr.)
44. Yamamoto, H., Aikawa, T., Matsutaka, H., Okuda, T., Ishikawa, I. 1974. *Am. J. Physiol.* 226:1428–33
45. Owen, E. E., Robinson, R. R. 1963. *J. Clin. Invest.* 42:263–75
46. Goldberg, A. L., Odessey, R. 1972. *Am. J. Physiol.* 223:1384–91
47. Ryan, N. T. et al 1974. *Fed. Proc.* 33:386
48. Odessey, R. 1973. *Regulation of amino acid metabolism in rat skeletal muscle.* PhD thesis, Harvard University
49. Jefferson, L. S., Rannels, D. E., Munger, B. L., Morgan, H. E. 1974. *Fed. Proc.* 33:1098–1104
50. Carlsten, A., Hallgren, B., Svanborg, A., Werko, L. 1967. *Acta Med. Scand.* 181:199–207
51. Cahill, G. F. Jr., Aoki, T. T., Marliss, E. B. 1972. *Handbook of Physiology. The Endocrine Pancreas,* ed. D. Steiner, N. Frienkel, 563–77. Baltimore: Williams and Wilkins
52. Goldberg, A. L. 1972. *Muscle Biology,* ed. R. G. Cassens, 1:89–118. New York: Dekker
53. Millward, D. J. 1970. *Clin. Sci.* 39:577–603
54. Garber, A. J. Karl, I. E., Kipnis, D. M. 1974. *Clin. Res.* 22:468A
55. London, D. R., Prenton, M. A. 1968. *Clin. Sci.* 35:55–61
56. Cahill, G. F. Jr. 1971. *The Human Adrenal Cortex,* ed. N. P. Christy, 205–38. New York: Harper & Row
57. Owen, O. E., Cahill, G. F. Jr. 1973. *J. Clin. Invest.* 52:2596–2605
58. Young, V. R., Haverburg, L. N., Bilmazes, C., Munro, H. N. 1973. *Metabolism* 22:1429–36
59. Cahill, G. F., Aoki, T. T., Ruderman, N. B. 1972. *Trans. Am. Clin. Climatol. Assoc.* 184–202

60. Carlsten, A., Hallgren, B., Jagenburg, R., Svanborg, A., Werko, L. 1961. *Scand. J. Clin. Lab. Invest.* 13:418–28
61. Felig, P., Marliss, E., Owen, O. E., Cahill, G. F. Jr. 1969. *Advan. Enzyme Regul.* 7:41–46
62. Exton, J. H. et al 1970. *Recent Progr. Horm. Res.* 26:411–57
63. Pagliara, A. S., Karl, I. E., DeVivo, D. C., Feigen, R. D., Kipnis, D. M. 1972. *J. Clin. Invest.* 51:1440–49
64. Herrera, E., Knopp, R. H., Freinkel, N. 1969. *J. Clin. Invest.* 48:2260–72
65. Felig, P., Lynch, V. 1970. *Science* 170:990–92
66. Felig, P., Kim, Y. J., Lynch, V., Hendler, R. 1972. *J. Clin. Invest.* 51:1195–1202
67. Ruderman, N. B., MacDonald, M., Neufeldt, N., Park, B., Berger, M. 1974. *Liver Metabolism.* Benzon Foundation Symposium VI, ed. F. Lundquist, N. Tygstrup, 291–300. Copenhagen: Munksgaard
68. Arroyave, G., Wilson, D., DeFures, C., Behar, M. 1962. *Am. J. Clin. Nutr.* 11:517–24
69. Holt, L. E., Snyderman, S. E., Norton, P. M., Roitman, E., Finch, J. 1963. *Lancet* 2:1343–48
70. Slone, D., Taitz, L. S., Gilchrist, G. D. 1961. *Brit. Med. J.* 1:32–34
71. Silbert, C. K. et al 1974. *Clin. Res.* 22:479A (Abstr.)
72. Carlsten, A., Hallgren, B., Jagenburg, R., Svanborg, A., Werko, L. 1966. *Acta Med. Scand.* 179:361–70
73. Felig, P., Marliss, E., Ohman, J. L., Cahill, G. F. Jr. 1970. *Diabetes* 19:727–29
74. Hambreus, L., Westphal, O., Hagberg, G. 1972. *Acta Paediat. Scand.* 61:81–88
75. Long, C. L., Spencer, J. L., Kinney, J. M., Geiger, J. W. 1971. *J. Appl. Physiol.* 31:110–16
76. Cuthbertson, D. P. 1964. *Mammalian Protein Metabolism,* ed. H. N. Munro, J. B. Allison, 373–414. New York: Academic
77. Allison, S. P., Hinton, P., Chamberlain, M. J. 1968. *Lancet* 2:1113–16
78. Hinton, P., Allison, S. P., Littlejohn, S., Lloyd, J. 1971. *Lancet* 1:767–69
79. Rubin, D. S., Hochgelerent, E., Rubin, A. L., Stenzel, K. H. 1972. *Lancet* 2:34–37
80. Giovannetti, S., Balestri, P. C., Biagini, M., Menichini, G., Rindi, P. 1970. *Arch. Intern. Med.* 126:900–5
81. Richards, P., Brown, C. L., Houghton, B. J., Thompson, E. 1971. *Lancet* 2:128–34
82. Walser, M., Coulter, A. W., Dighe, S., Crantz, F. R. 1973. *J. Clin. Invest.* 52:678–90
83. Close, J. H. 1974. *N. Engl. J. Med.* 290:663–67
84. Rudman, D. 1971. *J. Clin. Invest.* 50:90–96
85. Walser, M., Lund, P., Ruderman, N. B., Coulter, A. W. 1973. *J. Clin. Invest.* 52:2865–77

Copyright 1975. All rights reserved

LOW RENIN HYPERTENSION ❖7148

J. Caulie Gunnells, Jr., M.D. and William L. McGuffin, Jr., M.D.[1]
Division of Nephrology, Department of Medicine, Duke University School of Medicine, Durham, North Carolina 27710

Low or suppressed plasma renin activity was first described in association with primary aldosteronism in 1964 by Conn et al (1). They suggested that low plasma renin activity (PRA) was an essential component of the diagnostic criteria for primary aldosteronism and that renin measurements were useful in the evaluation of hypertensive patients in the search for the diagnosis of primary aldosteronism. Based on the finding of low or suppressed PRA in approximately 20 to 25% of their hypertensive patients, together with a previously reported high incidence of adrenal abnormalities in postmortem studies of patients with hypertension, they suggested that a normokalemic form of primary aldosteronism might comprise as many as 20% of patients otherwise classified as having benign essential hypertension (2). In the following year they described a patient with hypertension, suppressed PRA, normal serum potassium levels, and an elevated aldosterone excretion (3). In the several years following these initial reports, numerous other investigators measured PRA in large numbers of hypertensive patients, selected and unselected, finding low or suppressed levels of PRA in 10–43%; however, in only a relatively small number of these patients was the diagnosis of primary aldosteronism actually established (4–22). These findings led several investigators to suggest that low plasma renin activity was a nonspecific finding which was unrelated to abnormalities of adrenal structure or function (4, 23).

Between 1969 and 1971, three groups of investigators made preliminary observations suggesting that low renin hypertension was different from the usual variety of benign essential hypertension, and was related pathogenetically to the adrenal gland (16, 19, 22). In 1969, Woods et al (22) documented that aminoglutethimide, an inhibitor of adrenal steroidogenesis, produced a reduction of blood pressure in six of nine patients with suppressed renin activity and normal or low aldosterone levels. The second observation came from our own laboratories (16) in which four of six patients with low PRA and normal aldosterone excretion rates exhibited a reduction in blood pressure following subtotal adrenalectomy. Five of those patients had

---

[1] This work supported in part by Training Grant HL05848, Renal Division and CRU Grant MOIRR30.

pathological abnormalities in the adrenal including one patient with a single adenoma. Moreover, postoperative studies in these patients demonstrated a return of renin responsiveness. The third set of observations came from Spark & Melby in 1971 (19), in which eight patients with low PRA and low aldosterone secretion rates responded to high dose (400 mg/day) spironolactone therapy with a normalization of blood pressure and significant weight loss. This was in contrast to 15 patients who had normal PRA who failed to show an appreciable change in blood pressure or weight reduction using similar doses of spironolactone over a five week period.

It has become clear in the several years since these observations that hypertension associated with low PRA is a spectrum of disorders including primary aldosteronism with either adenoma or hyperplasia, various other forms of hypermineralocorticoidism, low renin essential hypertension with no easily definable abnormality in mineralocorticoid production, but in which the hypertension seems clearly to be volume-dependent, and several other miscellaneous conditions.

The observations cited above have led to considerable controversy regarding the criteria for establishing the low renin state, pathophysiologic mechanisms, as well as the indications and modalities for treatment, and prognosis. The purpose of this paper is to review recent developments that have been the source of controversy in the syndrome of low renin hypertension.

DEFINITION OF LOW RENIN HYPERTENSION

Low renin hypertension is characterized by the presence of hypertension in association with a failure to demonstrate an appropriate rise in PRA in response to stimuli (posture, sodium restriction, and pharmacologic agents) that consistently stimulate a rise in PRA in other hypertensive patients or normal subjects (6, 8, 9, 11–14, 17–20, 24–28). Most investigators have recognized the difficulty of distinguishing low levels of renin activity from normal levels of renin activity in hypertensive patients on an ad lib sodium intake.

The results of an investigation of renin responsiveness in 134 hypertensive patients from our own laboratory (Figure 1 and Table 1) demonstrate a 26% incidence of suppressed PRA. In this study a five day period of sodium restriction (10 meq/day) plus diuretics on the fourth day was used to stimulate PRA. These results are comparable to those reported by others using a variety of maneuvers known to provoke an increase in PRA. The limitations of this somewhat lengthy approach to stimulate renin release have been replaced in part by the outpatient administration of "loop" diuretics such as furosemide either orally or intravenously, and the failure of patients to respond to such maneuvers with an appropriate increase in PRA has been judged to be evidence for the low renin state (29). In those studies in which the latter approach has been compared with previous protocols utilizing dietary sodium restriction, there has been generally good agreement and it is now accepted by most investigators that the use of furosemide to demonstrate the low renin state is a valid diagnostic maneuver. However, each group of investigators has used slightly different protocols to elevate the PRA, making an evaluation of the literature difficult.

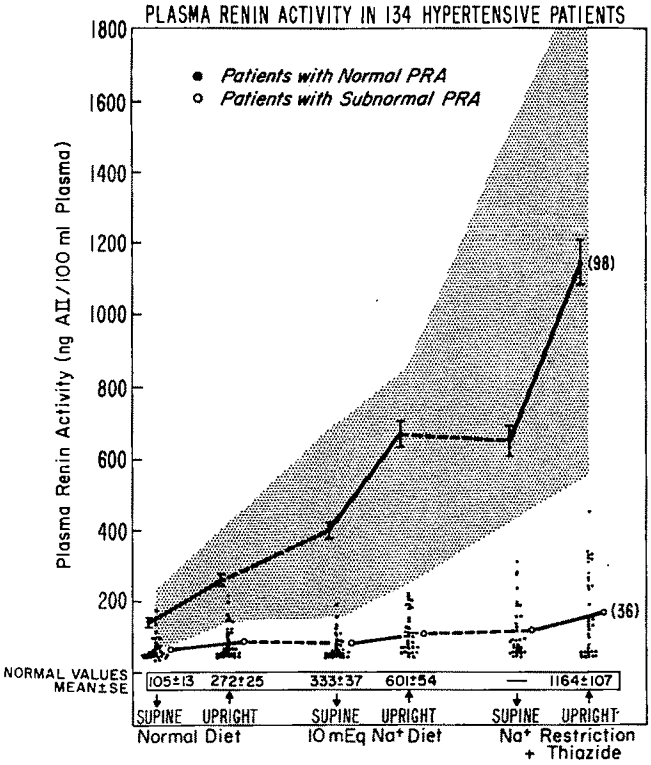


Figure 1 The results of plasma renin activity (PRA) measurements in 134 hypertensive patients on a normal diet and following dietary sodium restriction alone and the subsequent administration of diuretics. The shaded area corresponds to the range of PRA values obtained in normal subjects on a similar protocol; the upper line reflects the average PRA of the 98 hypertensive subjects who responded to these maneuvers with an appropriate increase in PRA as compared to normal subjects. The individual dots and the lower line (mean value) represent the PRA results in the 36 patients (26%) who were judged to have suppressed or subnormal PRA defined, in this study as a failure to increase PRA in the upright position on the 4th day of Na^+ restriction alone or on the 5th day after the administration of diuretic to at least the lowest level of PRA observed in normal subjects.

The necessity to use a stimulatory maneuver or a lengthy protocol would, on the surface, seem to have been solved by Laragh's group (30), which has related PRA and aldosterone secretion rates to the 24-hr urinary sodium excretion, thereby dividing patients into low, normal, and high renin subgroups. Originally they stated that these renin subgroups were valid at any level of urinary sodium excretion; however, in more recent publications (31–33), they have not been able to clearly distinguish some patients with normal renin activity from other patients with low renin activity at high levels of sodium excretion and have therefore modified their

Table 1 The mean plasma renin activity (PRA) values in 134 hypertensive subjects (see Figure 1) studied under standardized protocol conditions to detect low or suppressed PRA.[a]

Subjects	Normal Diet[b]		10 meq Na^+ Diet[c]		10 meq Na^+ Diet[d] plus Diuretic
	Supine	Upright	Supine	Upright	Upright
Normotensive	105	272	333	601	1164
Hypertensive					
Normal PRA	141	261	398	667	1145
Low or Subnormal PRA	60	82	63	114	133

[a] Number of subjects: Normotensive = 16; Hypertensive = 134 (98 with normal PRA; 36 with low PRA).
[b] Normal diet = 180 meq Na^+ and at least 100 meq K^+.
[c] Plasma obtained on day 4 of 10 meq Na^+diet
[d] Oral hydrochlorothiazide, 50 mg × 2 on day 4; plasma obtained on day 5.

Table 2 Conditions associated with increased or decreased renin release

Increased Renin Release	Decreased Renin Release
Upright posture	Supine posture
Na^+ restriction (volume depletion)	Na loading (volume expansion)
K^+ depletion	K^+ administration
Stimulation of autonomic nervous system	Blockade of autonomic nervous system
Drugs:	Drugs:
Diuretics including spironolactone	Propranolol
Diazoxide	Methyldopa
Nitroprusside	Reserpine
Hydralazine	Ganglionic blockers
	Clonidine

criteria so that urinary sodium excretion at the time of renin measurements is limited to less than 150 meq/24 hr. At least one other group (34) of investigators has been able to apply similar criteria in their hypertensive patients and has obtained a similar division into low, normal, and high renin groups.

It is important to recognize some of the currently known physiologic variables which may effect levels of renin activity in general. One of these is the observation that PRA and aldosterone decrease with advancing age (35–37); that in general PRA is lower in females as compared to males (20, 34); and in most studies PRA is lower in blacks (7, 13, 30, 34) than in whites and other racial groups. Therefore, in evaluating a hypertensive patient with PRA measurements, one must compare the values obtained to the PRA in age-, sex-, and race-matched controls. This may not be a major source of error, but is one of which the physician must be aware.

It is also important for the clinician to recognize the large number of other variables that may alter renin release if he is to accurately evaluate the results of PRA measurements. For a full review of renin physiology and its feedback control, the reader is referred to two recent publications (28, 38). Some commonly recognized conditions associated with increased or decreased renin release are outlined in Table 2. Careful attention must be paid to the status of sodium and potassium balance (39), posture, current and past drug history, to mention a few, if one is to avoid pitfalls in the renin profiling of hypertensive patients. Many of the drugs used in hypertensive therapy inhibit renin secretion including propranolol (40, 41) and other adrenergic blockers (42, 43), methyldopa (44), reserpine, ganglionic blockers (45), and clonidine (46, 47). Others may stimulate renin secretion including spironolactone (48), other diuretics (16, 29, 49), several vasodilators such as diazoxide (6, 50), nitroprusside (27), and hydralazine (51). These drugs should be withdrawn for at least a two week period prior to study.

Moreover, of equal importance, is some knowledge about the actual methodological techniques employed by the laboratory performing the renin measurements. It should be demonstrated that the method employed is of sufficient sensitivity to detect the low levels of renin activity known to exist in patients with low renin hypertension (52, 53).

DIFFERENTIAL DIAGNOSIS OF LOW RENIN HYPERTENSION

The differential diagnosis of low renin hypertension can be divided into four broad categories as outlined in Table 3. These four categories are (*a*) elevated levels of aldosterone, (*b*) elevation in other mineralocorticoids, (*c*) low renin essential hyper-

Table 3 Differential diagnosis of low renin hypertension

I. Aldosterone Excess
    A. Aldosterone-producing adenoma (APA)
    B. Idiopathic hyperaldosteronism (IHA)
    C. Indeterminate aldosteronism (Ind A)
    D. Glucocorticoid suppressible aldosteronism

II. Mineralocorticoid Excess
    A. Congenital adrenal hyperplasia
        1. 11-β-Hydroxylase deficiency
        2. 17-α-Hydroxylase deficiency
    B. Ectopic ACTH syndrome
    C. Adrenal carcinoma
    D. Isolated DOC excess
    E. 18-OH-DOC excess

III. Low Renin Essential Hypertension

IV. Miscellaneous Conditions
    A. Licorice ingestion
    B. Liddle's syndrome
    C. Parenchymal renal disease (occasionally)

tension, and (*d*) a miscellaneous group of hypertensive patients with a variety of unrelated abnormalities.

ALDOSTERONE EXCESS

Elevated levels of aldosterone either in plasma or urine have been found in from 3 to 50% of patients with hypertension and low PRA. In most series, approximately 5–10% of patients with low renin activity have elevated levels of aldosterone giving an incidence of hyperaldosteronism in the overall hypertensive population of < 2%. Patients with low renin hypertension in association with elevated levels of aldosteronism can be classified into four subgroups (54 55), including: (*a*) aldosterone-producing adenomas (APA), (*b*) idiopathic alcosteronism (adrenal hyperplasia, IHA), (*c*) indeterminate aldosteronism, and (*d*) glucocorticoid suppressible aldosteronism. The differential diagnosis between these four categories is critically important because surgery with removal of an adrenal adenoma is curative in most patients who have aldosterone-producing adenomas (56–60); however, adrenalectomy in those patients with bilateral adrenal hyperplasia which is found in the remaining three types of aldosteronism does not, in most instances, correct the hypertension unless a total adrenalectomy is performed.

Patients with aldosterone-producing adenomas in general have lower serum potassium levels, a greater elevation in aldosterone secretion, and lower PRA levels than patients in the other groups (61). However, in individual patients, these findings do not allow for clinical separation between the other forms of aldosterone excess. An initial step in the differential diagnosis of these conditions is to determine that the excessive aldosterone production is nonsuppressible. This may be accomplished by volume expanding the patient either with saline loading (62–65), the administration of deoxycorticosterone acetate (DOCA) (54), or fludrocortisone (55, 66). Those patients in the indeterminate category will have a suppression of their aldosterone excess to normal. Patients in this group also will exhibit some elevation in their PRA and plasma aldosterone levels with upright posture. PRA is easily elevated following one month of low dose spironolactone therapy (100 mg/day), and is associated with a fall in blood pressure (67).

In contrast, patients with idiopathic aldosteronism (adrenal hyperplasia) and those with aldosterone-producing adenoma will have no suppression of aldosterone levels with these maneuvers. However, idiopathic aldosteronism may exhibit an increase in PRA and plasma aldosterone levels with upright posture, whereas those patients with adenomas fail to demonstrate an increase in PRA and aldosterone levels by this maneuver and, paradoxically, plasma aldosterone may decrease (68). Another differential point between these two forms of aldosteronism is that with moderate dose spironolactone therapy (200 mg/day) for one month, PRA will usually increase in patients with hyperplasia, but will not increase in patients with classic Conn's syndrome (67); however, with 400 mg/day the blood pressure of both will be decreased (69). Recently, one group has used quadric analysis (70, 71) of eight variables to arrive at the pre-operative distinction between patients with adenomas versus hyperplasia. The variables used were age, systolic and diastolic blood pressures, renin concentration, plasma aldosterone, and serum sodium, potassium, and bicarbonate concentrations. Ultimately, however, the differential between adrenal hyperplasia and adenoma is established by actually identifying the adenoma with either adrenal venography (72) or iodocholesterol scanning (73) of the adrenal gland in conjunction with the identification of differential levels of plasma aldosterone from adrenal vein effluent (74, 75).

The glucocorticoid suppressible form of hyperaldosteronism is extremely rare, with only six cases having been reported to date (76–78), and five of six patients were children with two each from the same family. These patients also usually have elevated levels of 18-hydroxycorticosterone or corticosterone. No enzymatic defect has been established in these patients to date. The hypertension and hypokalemia is easily corrected by low dose dexamethasone therapy.

MINERALOCORTICOID EXCESS

Other forms of mineralocorticoid hypertension (Table 3) are being discovered or described with increasing frequency as methods for measuring these hormones become more widespread. Nonaldosterone mineralocorticoids, capable of producing sodium retention and volume expansion, which have been associated with hypertension, are deoxycorticosterone (DOC), corticosterone (B), and 18-hydroxydeoxycorticosterone (18-OH-DOC).

The first category represents enzymatic defects of adrenal steroidogenesis manifested by congenital adrenal hyperplasia (79) involving either 17-α-hydroxylase or 11-β-hydroxylase deficiency. 17-α-hydroxylation deficiency (80–83) results in elevated levels of DOC, B, and 18-hydroxycorticosterone. The distinguishing clinical features of this syndrome are primary amenorrhea in the female and male pseudohermaphroditism. The second variety of congenital adrenal hyperplasia producing hypertension is the 11-β-hydroxylation deficiency syndrome (84, 85) recognized by virilization in infants and children and the frequent presence of hypokalemia. In this syndrome the secretion of both 11-deoxycortisol and DOC are increased with normal or reduced levels of aldosterone, B, and cortisol. In both forms of congenital adrenal hyperplasia, suppression of ACTH with glucocorticoids will correct the hypertension as well as the electrolyte abnormalities.

Elevated levels of DOC and/or compound B have been reported in association with hypertension in the absence of definable enzymatic defects (85–88). Elevations in these two compounds have been reported in patients with ectopic ACTH syndromes and in patients with adrenal cortical carcinoma. It has been observed that patients with adrenal cortical carcinoma may have elevated levels of aldosterone (89) as well as DOC or B. Recently, elevated levels of DOC alone have been reported in 6 of 21 patients with low renin hypertension (90, 91). These patients had no evidence of neoplasia. The hypertension in these patients was corrected with spironolactone; however, dexamethasone treatment was ineffective, suggesting that the elevated levels of DOC were not ACTH-dependent.

Recently, attention has been directed to a relatively low potency human mineralocorticoid, 18-OH-DOC, first reported to be elevated in 3 of 12 low renin hypertensive patients by Melby et al (92). In one patient treated with dexamethasone blood pressure returned to normal, suggesting that this mineralocorticoid was under

ACTH control. They have since described elevated levels of this compound in 10 of 30 patients with low renin hypertension (93). The adrenal abnormalities in most of these patients were hyperplasia, although two had adrenal adenomas. In those patients with hyperplasia, dexamethasone reduced the blood pressure and adrenalectomy was likewise effective in normalizing the pressure. They also stated that approximately 15% of a random selection of low renin patients will have marked elevation of this mineralocorticoid. In addition, they have described elevated levels of this steroid in patients with otherwise classical forms of primary aldosteronism. It is of interest that Genest and co-workers (94) have demonstrated that over 60% of their patients with essential hypertension have elevated levels of 18-OH-DOC. The significance of these findings remains uncertain at the present time.

LOW RENIN ESSENTIAL HYPERTENSION

By far, the largest category of patients with low renin hypertension have no identifiable abnormality in mineralocorticoid secretion. While the precise pathophysiology in this group of hypertensive patients has not been definitively established, certain observations have been made which suggest either mild abnormalities in mineralocorticoid control or renal handling of sodium such that a state of effective volume expansion exists in these patients.

Subtle abnormalities of aldosterone secretion have been documented both in low renin hypertensive patients and in benign essential hypertension. Studies from Luetscher's laboratory have revealed that aldosterone secretion could not be suppressed with sodium loading in both normal and low renin hypertensive patients (17). In those patients with normal plasma renin activity, aldosterone secretion could be stimulated with a low sodium diet; however, in patients with low renin levels, aldosterone secretion was unchanged by this maneuver. Similar studies have been reported by others (18, 20, 95). Higher levels of aldosterone per unit of PRA have been reported in low renin hypertensive patients as compared to normal renin hypertensives or to normal subjects (96–99). This finding suggests that the adrenal cortex in patients with low renin hypertension is more sensitive to stimulation by angiotensin II than is the adrenal cortex in normal renin hypertensives. The low salivary sodium to potassium concentrations that have been described in low renin hypertensive patients (100) have also been interpreted as evidence supporting the existence of a mineralocorticoid effect.

Additional data for an abnormality in mineralocorticoid control in low renin hypertension come from studies using spironolactone. It was initially suggested that spironolactone, an inhibitor of mineralocorticoid activity, was a specific form of therapy for this variety of hypertension (48, 101, 102), and that other diuretics were less effective, whereas spironolactone alone was relatively ineffective in patients with normal renin hypertension (19). More recent evidence has suggested that thiazide diuretics effectively lower blood pressure in patients with low renin essential hypertension (103–105); however, the blood pressure response to thiazide diuretics is less than with spironolactone at comparable dose levels (106). At least one group of investigators (103) has demonstrated no difference in the blood pressure response to these two diuretics in patients with low renin hypertension.

Since in general, spironolactone has been shown to be relatively more effective than thiazide diuretics in patients with low renin hypertension, this may be used as suggestive evidence for a mineralocorticoid-mediated hypertensive mechanism; however, it certainly does not establish that point. The previously mentioned studies demonstrating the effectiveness of aminoglutethimide in these patients, together with the observed alterations in adrenal pathology, offer additional evidence for the participation of the adrenal cortex in the pathogenesis of the hypertension in this group of patients.

While it has been clearly established that patients with primary aldosteronism have increased plasma and ECF volumes, as well as increased exchangeable sodium space in comparison with both normal subjects and patients with essential hypertension (107–110), conclusive evidence of similar abnormalities in low renin hypertension has not been presented.

Helmer & Judson (7) found a small but insignificant increase in blood volume in low renin hypertensive patients and similar results were presented by Jose et al (15). However, there was no difference in blood volumes between the low and normal renin hypertensive patients in studies by Birkenhäger (111) and Woods et al (22). A statistically significant increase in mean ECF volume in low renin hypertensives has been recently reported from Kaplan's group (15). In addition, Woods (22) found a significantly greater exchangeable sodium space in low renin hypertensive patients as compared to normal renin hypertensive states; however, there was no difference between the low renin activity group and their matched normotensive controls. More recently, another group has found no difference in exchangeable sodium space between low and normal renin hypertensive patients (110).

While it is clear that volume expansion comparable to that found in primary aldosteronism is not present in low renin hypertension, the evidence does suggest a mild abnormality of volume expansion in many of these patients. Perhaps the most convincing argument for volume excess being a primary pathophysiological derangement in low renin hypertensive patients is their uniform response to diuretic administration with significant weight loss and reduction in blood pressure as compared to normal renin hypertensive patients.

MISCELLANEOUS CONDITIONS

There exists a small group of patients with low renin hypertension who have a variety of miscellaneous abnormalities. These include those patients with excessive licorice ingestion (112, 113) in whom hypokalemia and hypertension develop and may mimic primary aldosteronism, but in whom the abnormal compound has been identified as glycyrrhizinic acid.

Liddle and co-workers (114) reported in 1963 on a patient with a renal tubular defect leading to potassium wasting and sodium retention in association with low

PRA levels. These patients failed to respond to spironolactone but responded to triamterene, suggesting a non-mineralocorticoid mechanism for their potassium wasting.

There are some patients with parenchymal renal disease and renal insufficiency in association with hypertension who have low PRA levels. A volume-dependent form of hypertension has been postulated based on normalization of blood pressure with ultrafiltration on hemodialysis, whereas patients who fail to respond to ultrafiltration have demonstrated high renin activity and have required renin blocking agents or nephrectomy for control of hypertension.

TREATMENT

The first line of treatment in patients with low renin essential hypertension clearly involves the use of diuretic agents which in most of these patients reduce the blood pressure to normal. As mentioned above, relatively greater success has been reported with the use of spironolactone than with the use of thiazide diuretics alone. More recent studies have demonstrated that many patients with low renin essential hypertension respond very well to the thiazide diuretics, and the response to spironolactone can no longer be utilized as a unique feature of the low renin hypertensive state. However, spironolactone has been particularly useful in the medical management of low renin hypertension with aldosterone excess (Conn's syndrome). The use of spironolactone in all forms of low renin hypertension avoids the well known kaliuretic effect of thiazides, which may be enhanced in states of mineralocorticoid excess, thereby minimizing the need for supplemental potassium. Potential negative features of spironolactone are the associated gastrointestinal disturbances such as persistent nausea, as well as menorrhagia and gynecomastia, which may occur with its administration. In the low renin essential hypertensives this is largely a theoretical objection since the dosages that have been effective are generally lower (100–200 mg/day) and have a low incidence of these side effects in comparison with the earlier recommended high dose therapy (400 mg/day). If diuretic agents alone are not successful in reducing blood pressure to normal, then standard antihypertensive drugs are effective in this group of patients and should be used. The currently recognizable clinical entities associated with low renin activ-

Table 4 A treatment scheme for low renin hypertension (see text)

Diagnosis	Low Renin Activity		Treatment		
	Aldosterone Increased	Aldosterone Normal or Decreased	Surgical	Medical	
				Steroids[a]	Diuretic[b]
Conn's syndrome	X		X		X
Idiopathic hyperaldosteronism	X				X
17-α-Hydroxylase defect		X		X	
11-β-Hydroxylase defect		X		X	
18-OH-DOC		X		X	X
DOC excess		X			X
Low renin "essential"		X			X

[a] Corticosteroids.

[b] Spironolactone vs thiazide diuretics.

ity together with their mineralocorticoid defects and a proposed scheme for treatment are outlined in Table 4.

PROGNOSIS OF PATIENTS WITH LOW RENIN HYPERTENSION

Perhaps the greatest area of controversy in this syndrome has been generated by the report of Brunner et al (30) that patients with low renin hypertension had fewer cerebral vascular accidents and myocardial infarctions than patients with normal or high renin hypertension. They originally reported on 59 patients with low renin hypertension in which they found no strokes or myocardial infarctions as contrasted with an 11% and 14% incidence in the normal and high PRA subgroups respectively. Recently Christlieb et al (115) have reported data supporting this concept. The conclusions of Laragh's group relative to the benign nature of low renin hypertension is at variance with the more recently reported data from five other laboratories (34, 116–119). The results of these five studies clearly indicate that patients with low renin hypertension have an equal number of morbid cardiovascular events, that is, myocardial infarction, cerebral vascular accidents, and congestive heart failure, as do patients with normal or high renin hypertension. Laragh has contended that these investigators are studying a different group of hypertensive patients because they have not followed a similar protocol of renin profiling. This argument is not valid particularly in view of the study from Catt's laboratory (34) where similar procedures were followed as in Laragh's study with the exception of minor variations in the technique of renin activity measurements. They demonstrated a similar incidence of low, normal, and high renin hypertensives in their population, albeit black, and documented an equal incidence of myocardial infarctions and cerebral vascular accidents in the three renin subgroups. Furthermore, Genest (119) and co-workers have demonstrated that the risk of myocardial infarction or cerebral vascular accident is more closely related to the level of diastolic blood pressure, level of renal function, and efficacy of control of the hypertension than it is to the level of renin activity. Similar conclusions were reached by Gulati et al (117) who demonstrated an actual increase in the number of morbid events in their group of patients with low renin hypertension; however, in their study the levels of blood pressure in patients with low renin hypertension were significantly higher than in those with normal or high renin hypertension.

The results of the VA cooperative study (120, 121) on hypertension in which a large number of unselected hypertensives were followed for five years or more without reference to renin activity levels have demonstrated the beneficial effect of blood pressure reduction in decreasing the number of morbid events in effectively treated hypertensives. The evidence, therefore, seems conclusive that low renin hypertension is not necessarily a benign form of hypertension, is not different in terms of its predisposition to myocardial infarction, cerebral vascular accidents, and congestive heart failure from the other forms of hypertension, and it should not be ignored. It is clear that this hypertensive syndrome must be treated as vigorously as other forms of hypertension if one is to reduce the risk of morbid cardiovascular events.

SUMMARY

Low renin hypertension comprises a spectrum of disorders ranging from primary aldosteronism to obvious disorders of other mineralocorticoids, as well as a variety of miscellaneous disorders. The largest group of patients with low renin hypertension have no clear abnormality in mineralocorticoid production. However, many lines of evidence suggest the critical role of volume excess in the pathogenesis of hypertension in these patients. More detailed physiological studies must be performed in order to totally understand the spectrum of pathophysiology in low renin hypertension. However, while such studies are in progress the evidence from the literature suggests that these patients must be treated and that in most of these patients diuretic adminstration, either spironolactone or the thiazide group of diuretics, is usually effective in achieving a reduction of blood pressure to normal.

ACKNOWLEDGMENT

The authors acknowledge with grateful appreciation the expert secretarial assistance of Mrs. Iris Pritchard.

Literature Cited

1. Conn, J. W., Cohen, E. L., Rovner, D. R. 1964. Suppression of plasma renin activity in primary aldosteronism: Distinguishing primary from secondary aldosteronism in hypertensive disease. *J. Am. Med. Assoc.* 190:213–21
2. Conn, J. W. 1964. Plasma renin activity in primary aldosteronism: Importance in differential diagnosis and in research of "essential" hypertension. *J. Am. Med. Assoc.* 190:222–25
3. Conn, J. W., Cohen, E. L., Rovner, D. R., Nesbit, R. M. 1965. Normokalemic primary aldosteronism: A detectable cause of curable "essential" hypertension. *J. Am. Med. Assoc.* 193:100–6
4. Ledingham, J. G. G., Bull, M. B., Laragh, J. H. 1967. The meaning of aldosteronism in hypertensive disease. *Circ. Res.* 20 & 21 (Suppl. 2):177–86
5. Creditor, M. C., Loschky, U. K. 1968. Incidence of suppressed renin activity and of normokalemic primary aldosteronism in hypertensive negro patients. *Circulation* 38:1027–31
6. Kuchel, O., Fishman, L. M., Liddle, G. W., Michelakis, A. 1967. Effect of diazoxide on plasma renin activity in hypertensive patients. *Ann. Intern. Med.* 67:791–99
7. Helmer, O. M., Judson, W. E. 1968. Metabolic studies on hypertensive patients with suppressed plasma renin activity not due to hyperaldosteronism. *Circulation* 38:965–76
8. Gunnells, J. C., Grim, C. E., Robinson, R. R., Wilderman, N. M. 1967. Plasma renin activity in healthy subjects and patients with hypertension. *Arch Intern. Med.* 119:232–40
9. Weinberger, M. H., Dowdy, A. J., Nokes, G. W., Leutscher, J. A. 1968. Plasma renin activity and aldosterone secretion in hypertensive patients during high and low sodium intake and administration of diuretic. *J. Clin. Endocrinol. Metab.* 28: 359–71
10. Fishman, L. M. et al 1968. Incidence of primary aldosteronism in uncomplicated essential hypertension. *J. Am. Med. Assoc.* 205:497–502
11. Kaneko, Y. et al 1968. Renin release in patients with benign essential hypertension. *Circulation* 38:353–62
12. Creditor, M. C., Loschky, U. K. 1967. Plasma renin activity in hypertension. *Am. J. Med.* 43:371–81
13. Channick, B. J., Adlin, E. V., Marks, A. D 1969. Suppressed plasma renin activity in hypertension. *Arch. Intern. Med.* 123:131–40
14. Jose, A., Kaplan, N. M. 1969. Plasma renin activity in the diagnosis of primary aldosteronism. *Arch Intern. Med.* 123: 141–46
15. Jose, A., Crout, J. R., Kaplan, N. M. 1970. Suppressed plasma renin activity

in essential hypertension. *Ann. Intern. Med.* 72:9–16
16. Gunnells, J. C. et al 1970. Hypertension, adrenal abnormalities, and alterations in plasma renin activity. *Ann. Intern. Med.* 73:901–11
17. Collins, R. D. et al 1970. Abnormally sustained aldosterone secretion during salt loading in patients with various forms of benign hypertension; relation to plasma renin activity. *J. Clin. Invest.* 49:1415–26
18. Kotchen, T. A., Murrow, P. J., Morrow L. B., Shutkin, P. M., Marieb, N. 1971. Renin and aldosterone in essential hypertension. *Clin. Sci.* 41:321–31
19. Spark, R. F., Melby, J. C. 1971. Hypertension and low plasma renin activity: presumptive evidence for mineralocorticoid excess. *Ann. Intern. Med.* 75:831–36
20. Crane, M. G., Harris, J. J., Johns, V. J. Jr. 1972. Hyporenemic hypertension. *Am. J. Med.* 52:457–66
21. Streeten, D. H. P., Schletter, F. E., Clift, G. V., Stevenson, C. T., Dalakos, T. G. 1969. Studies of the renin angiotensin-aldosterone system in patients with hypertension and in normal subjects. *Am. J. Med.* 46:844–61
22. Woods, J. W., Liddle, G. W., Stant, E. G. Jr., Michelakis, A. M., Brill, A. B. 1969. Effect of an adrenal inhibitor in hypertensive patients with suppressed renin. *Arch. Intern. Med.* 123:366–70
23. Kaplan, N. M. 1969. Commentary on incidence of primary aldosteronism: Current estimations based on objective data. *Arch. Intern. Med.* 123:152–54
24. Cohen, E. L., Rovner, D. R., Conn, J. R. 1966. Postural augmentation of plasma renin activity. *J. Am. Med. Assoc.* 197:973–78
25. Kem, D. C. et al 1972. The low renin state: definition and implications. *Clin. Res.* 20:776
26. Luetscher, J. A. et al 1969. Effects of sodium loading, sodium depletion and posture on plasma aldosterone concentration and renin activity in hypertensive patients. *J. Clin. Endocrinol. Metab.* 29:1310–18
27. Kameko, Y., Ikeda, T., Takeda, T., Ueda, H. 1967. Renin release during acute reduction of arterial pressure in normotensive subjects and patients with renovascular hypertension. *J. Clin. Invest.* 46:705–16
28. Stein, J. H., Ferris, T. F. 1973. The physiology of renin. *Arch. Intern. Med.* 131:860–72
29. Haber, E., Koerner, T., Page, L. B. Kliman, B., Purnode, A. 1969. Application of a radioimmunoassay for angiotensin I to the physiologic measurements of plasma renin activity in normal human subjects. *J. Clin. Endocrinol. Metab.* 29:1349–55
30. Brunner, H. R. et al 1972. Essential hypertension: renin and aldosterone, heart attack and stroke. *N. Engl. J. Med.* 286:441–49
31. Brunner, H. R., Sealey, J. E., Laragh, J. H. 1973. Renin as a risk factor in essential hypertension: more evidence. *Am. J. Med.* 55:295–302
32. Brunner, H. R., Sealey, J. E., Laragh, J. H. 1973. Renin subgroups in essential hypertension. Further analysis of their pathophysiological and epidemiological characteristics. *Circ. Res.* 32–33 (Suppl I):99–109
33. Laragh, J. H. et al 1972. Renin, angiotensin and aldosterone system in pathogenesis and management of hypertensive vascular disease. *Am. J. Med.* 52:633–52
34. Mroczek, W. J., Finnerty, F. A., Catt, K. J. 1973. Lack of association between plasma-renin and history of heart attack or stroke in patients with essential hypertension. *Lancet* ii:464–69
35. Tuck, M., Williams, G. H., Cain, J. P., Sullivan, J. M., Dluhy, R. G. 1973. Relation of age, diastolic pressure and known duration of hypertension to presence of low renin essential hypertension. *Am. J. Cardiol.* 32:637–42
36. Sambhi, M. P., Crane, M. G., Genest, J. 1973. Essential hypertension: New concepts about mechanisms. *Ann. Intern. Med.* 79:411–24
37. Schalekamp, M. A. D. H., Krauss, X. H., Schalekamp-Kuyken, M. P. A., Kolsters, G., Birkonhager, W. H. 1971. Studies on the mechanism of hypernatrunesis in essential hypertension in relation to measurements of plasma renin concentration, body fluid compartments and renal function. *Clin. Sci.* 41:219–31
38. Davis, J. O. 1973. The control of renin release. *Am. J. Med.* 55:333–50
39. Vander, A. J. 1970. Direct effects of potassium on renin secretion and renal function. *Am. J. Physiol.* 219:455–59
40. Buhler, F. R., Laragh, J. H., Baer, L., Vaughan, E. D., Brunner, H. R. 1972. Propranol inhibition of renin secretion. A specific approach to diagnosis and treatment of renin-dependent hyperten-

sive disease. *N. Engl. J. Med.* 287:1209–14

41. Stokes, G. S., Weber, M. A., Thornell, I. R. 1974. β-blockers and plasma renin activity in hypertension. *Brit. Med. J.* 1:60–62
42. Winer, N., Chokshi, D. S., Yoon, M. S. 1969. Adrenergic receptor mediation of renin secretion. *J. Clin. Endocrinol. Metab.* 29:1168–75
43. Michelakis, A. M., McAllister, R. G. 1972. The effect of chronic adrenergic receptor blockade on plasma renin activity in man. *J. Clin. Endocrinol. Metab.* 34:386–98
44. Mohammed, S. et al 1969. Effect on methyldopa on plasma renin activity in man. *Circulation* 35:593–98
45. Kameko, Y. et al 1970. Effect of ganglion-blocking agents on renin release in hypertensive patients. *Circ. Res.* 27:97
46. Baer, L., Brunner, H. R., Buhler, F. R., Laragh, J. H. 1971. Suppression of renin and aldosterone by clonidine. *Ann. Intern. Med.* 74:830
47. Onesti, G., Schwartz, A. B., Kim, K. E., Paz-Martinez, V., Swartz, C. 1971. Antihypertensive effect of clonidine. *Circ. Res.* 28–29 (Suppl. II):53–69
48. Spark, R. F., O'Hare, C. M. 1974. Low renin hypertension. Restoration of normotension and renin responsiveness. *Arch. Intern. Med.* 133:205–11
49. Bourgoignie, J. J., Catanzaro, F. J., Perry, H. M. Jr. 1968. Renin-angiotensin-aldosterone systems during chronic thiazide therapy of benign essential hypertension. *Circulation* 37:27–35
50. Baer, L., Goodwin, F. J., Laragh, J. H. 1969. Diazoxide-induced renin release in man: Dissociation from plasma and extracellular fluid volume changes. *J. Clin. Endocrinol. Metab.* 29:1107–9
51. Ueda, H., Yagi, S., Kaneko, U. 1968. Hydralazine and plasma renin activity. *Arch. Intern. Med.* 122:387–91
52. Sealey, J. E., Gerten-Banes, J., Laragh, J. H. 1972. The renin system: Variations in man measured by radioimmunoassay. *Kidney Int.* 1:240–53
53. Sealey, J. E., Laragh, J. M. 1973. Searching out low renin patients: Limitations of some commonly used methods. *Am. J. Med.* 55:303–14
54. Biglieri, E. G., Slaton, P. E. Jr., Kronfield, S. J., Schambelan, M. 1967. Diagnosis of an aldosterone-producing adenoma in primary aldosteronism: An evaluative maneuver. *J. Am. Med. Assoc.* 201:510–14
55. Biglieri, E. G., Stockigt, J. R., Schambelan, M. 1970. A preliminary evaluation for primary aldosteronism. *Arch. Intern. Med.* 126:1004–7
56. Biglieri, E. G., Schambelan, M., Slaton, P. E., Stockigt, J. R. 1970. The intercurrent hypertension of primary aldosteronism. *Circ. Res.* 26–27 (Suppl I):195–202
57. Brown, J. J. et al 1972. Comparison of surgery and prolonged spironolactone therapy in patients with hypertension, aldosterone excess, and low plasma renin. *Brit. Med. J.* 2:729–34
58. George, J. M., Wright, L., Bell, N. H., Bartter, F. C. 1970. The syndrome of primary aldosteronism. *Am. J. Med.* 48:343–56
59. Laragh, J. H., Ledingham, J. G. G., Sommers, S. C. 1967. Secondary aldosteronism and reduced plasma renin in hypertensive disease. *Trans. Assoc. Am. Physicians* 80:168–82
60. Baer, L., Sommers, S. C., Krakoff, L. R., Newton, M. A., Laragh, J. H. 1970. Pseudo-primary aldosteronism. An entity distinct from true primary aldosteronism. *Circ. Res.* 26–27 (Suppl I):203–20
61. Stockigt, J. R., Collins, R. D., Noakes, C. A., Biglieri, E. G. 1971. Significance of precise determination of basal plasma renin in primary aldosteronism. *J. Clin. Invest.* 50:89(a)
62. Espiner, E. A., Tucci, J. R., Jagger, P. L, Lauler, D. P. 1967. Effect of saline infusions on aldosterone secretion and electrolyte excretion in normal subjects and patients with primary aldosteronism. *N. Engl. J. Med.* 277:1–7
63. Espiner, E. A. et al 1971. The pattern of plasma renin activity and aldosterone secretion in normal and hypertensive subjects before and after saline infusions. *Am. J. Cardiol.* 27:585–94
64. Kem, D. C., Weinberger, M. H., Mayes, D. M., Nugent, C. A. 1971. Saline suppression of plasma aldosterone in hypertension. *Arch. Intern. Med.* 128:380–86
65. Christlieb, A. R. et al 1971. The pattern of electrolyte excretion of normal and hypertensive subjects before and after saline infusions. A simple electrolyte formula for the diagnosis of primary aldosteronism. *Am. J. Cardiol.* 27:595–609
66. Horton, R. 1969. Stimulation and suppression of aldosterone in plasma of normal man and in primary aldosteronism. *J. Clin. Invest.* 48:1230–36

67. Biglieri, E. G. 1973. Personal communication. Clinical Conference National Institute of Arthritis, Metabolic, and Digestive Diseases
68. Ganguly, A., Melada, G. A., Luetscher, J. A., Dowdy, J. 1973. Control of plasma aldosterone in primary aldosteronism: Distinction between adenoma and hyperplasia. *J. Clin. Endocrinol. Metab.* 37:765–75
69. Spark, R. F., Melby, J. C. 1968. Aldosteronism in hypertension. The spironolactone response test. *Ann. Intern. Med.* 69:685–91
70. Aitchison, J. et al 1971. Quadric analysis in the preoperative distinction between patients with and without adrenocortical tumors in hypertension with aldosterone excess and low plasma renin. *Am. Heart J.* 82:660–71
71. Ferris, J. B. et al 1970. Hypertension with aldosterone excess and low plasma renin: Preoperative distinction between patients with and without adrenocortical tumor. *Lancet* ii:995–1000
72. Nicols, G. L., Mihy, H. A., Modlinger, R. S., Gabkilove, J. L. 1972. Percutaneous adrenal venography. *Ann. Intern. Med.* 76:899–909
73. Conn, J. W. et al 1972. Photoscanning of tumors in primary aldosteronism. *Hypertension '72,* ed. J. Genest, E. Koiw, 299–312. Berlin: Springer. 617 pp.
74. Horton, R., Finck, E. 1972. Diagnosis and localization in primary aldosteronism. *Ann. Intern. Med.* 76:885–90
75. Scoggins, B. A., Oddie, C. J., Hare, W. S. C., Coghlan, J. P. 1972. Preoperative lateralization of aldosterone-producing tumors in primary aldosteronism. *Ann. Intern. Med.* 76:891–97
76. Salti, I. S., Stiefel, M., Ruse, J. L., Laidlaw, J. C. 1969. Non-tumorous "primary" aldosteronism: I. Type relieved by glucocorticoid (glucocorticoid-remedial) aldosteronism. *Can. Med. Assoc. J.* 101:1–10
77. Giebink, G. S., Gotlin, R. W., Biglieri, E. G., Katz, F. H. 1973. A kindred with familial glucocorticoid-suppressible aldosteronism. *J. Clin. Endocrinol. Metab.* 36:715–23
78. New, M. I., Siegal, E. J., Peterson, R. E. 1973. Dexamethasone-suppressible hyperaldosteronism. *J. Clin. Endocrinol. Metab.* 37:93–100
79. Bongiovanni, A. M., Root, A. W. 1963. The adrenogenital syndrome. *N. Engl. J. Med.* 268:1283–89, 1342–51, 1391–99
80. Biglieri, E. G., Herron, M. A., Brust, N. 1966. 17-α-hydroxylation deficiency in man. *J. Clin. Invest.* 45:1946–54
81. Goldsmith, O. D., Solomon, M., Horton, R. 1967. Hypogonadism and mineralocorticoid excess: the 17-hydroxylase deficiency syndrome. *N. Engl. J. Med.* 277:673–77
82. New, M. I. 1970. Male pseudo-hermaphroditism due to 17-α-hydroxylase deficiency. *J. Clin. Invest.* 49:1930–41
83. Bricaire, H. et al 1972. A new male pseudo-hermaphroditism associated with hypertension due to a block of 17-α-hydroxylation. *J. Clin. Endocrinol. Metab.* 35:69–72
84. New, M. I., Seaman, M. P. 1970. Secretion rates of cortisol and aldosterone precursors in various forms of congenital adrenal hyperplasia. *J. Clin. Endocrinol. Metab.* 30:361–71
85. Biglieri, E. G., Stockigt, J. R., Schambelan, M. 1972. Adrenal mineralocorticoids causing hypertension. *Am. J. Med.* 52:623–32
86. Crane, M. G., Harris, J. J. 1966. Desoxycorticosterone secretion rates in hyperadrenocorticism. *J. Clin. Endocrinol. Metab.* 26:1135–43
87. Biglieri, E. G., Slaton, P. E., Schambelan, M., Kronfield, S. J. 1968. Hypermineralocorticoidism. *Am. J. Med.* 45:170–75
88. Biglieri, E. G., Stockigt, J. R., Collins, R. D., Schambelan, M. 1972. Mineralocorticoid hypertension: Recognition and results of treatment. See Ref. 73, pp. 319–25
89. Filipecki, S. et al 1972. Carcinoma of the adrenal cortex with hyperaldosteronism. *J. Clin. Endocrinol. Metab.* 35:225–29
90. Brown, J. J. et al 1972. Apparently isolated excess deoxycorticosterone in hypertension. *Lancet* ii:243–47
91. Brown, J. J. et al 1972. A review of spironolactone therapy in patients with hypertension, aldosterone excess and low plasma renin. Increased desoxycorticosterone in hypertension—another variant of the mineralocorticoid excess syndrome. See Ref. 73, pp. 313–19
92. Melby, J. C., Dale, S. L., Wilson, T. E. 1971. 18-hydroxydeoxycorticosterone in human hypertension. *Circ. Res.* 28–29 (Suppl. II):143–50
93. Melby, J. C., Dale, S. L., Grekin, R. J., Gaunt, R., Wilson, T. E. 1972. 18-hydroxy-11-deoxycorticosterone (18-OH-DOC) secretion in experimental

and human hypertension. See Ref. 73, pp. 350–60

94. Genest, J., Nowaczynski, W., Kuchel, O., Sasaki, C. 1972. Plasma progesterone levels and 18-hydroxdeoxycorticosterone secretion rate in benign essential hypertension in humans. See Ref. 73, pp. 293–98
95. Williams, G. H. et al 1970. Abnormal responsiveness of the renin aldosterone system to acute stimulation in patients with essential hypertension. *Ann. Intern. Med.* 72:317–26
96. Grim, C., Peters, T. 1972. Low renin hypertension: A state of inappropriate secretion of aldosterone. *Clin. Res.* 20:427
97. Grim, C. E. 1973. Demonstration of elevated plasma aldosterone in low renin hypertension. *Clin. Res.* 21:493
98. Laragh, J. H., Sealey, J., Brunner, H. R. 1972. The control of aldosterone secretion in normal and hypertensive man: Abnormal renin-aldosterone patterns in low renin hypertension. *Am. J. Med.* 53:649–63
99. Buhler, F. R., Laragh, J. H., Sealey, J. E., Brunner, H. R. 1973. Plasma aldosterone-renin interrelationships in various forms of essential hypertension. *Am. J. Cardiol.* 32:554–61
100. Aldin, E. V., Channick, B. J., Marks, A. D. 1969. Salivary sodium-potassium ratio and plasma renin activity in hypertension. *Circulation* 39:685–92
101. Carey, R. M., Douglas, J. G., Schweikert, R., Liddle, G. W. 1972. The syndrome of essential hypertension and suppressed plasma renin activity. Normalization of blood pressure with spironolactone. *Arch. Intern. Med.* 130:849–54
102. Crane, M. G., Harris, J. J. 1970. Effect of spironolactone in hypertensive patients. *Am. J. Med. Sci.* 260:311–30
103. Adlin, E. V., Marks, A. D., Channick, F. J. 1972. Spironolactone and hydrochlorothiazide in essential hypertension. *Arch. Intern. Med.* 130:855–58
104. Vaughan, E. D. Jr et al 1973. Volume factor in low and normal renin essential hypertension. Treatment with either spironolactone or chlorthialidone. *Am. J. Cardiol.* 32:523–32
105. Douglas, J. G., Hollifield, J. W., Liddle, G. W. 1974. Treatment of low renin essential hypertension. Comparison of spironolactone and a hydrochlorothiazide triamterene combination. *J. Am. Med. Assoc.* 227:518–21
106. Beevers, D. G. et al 1973. The use of spironolactone in the diagnosis and the treatment of hypertension associated with mineralocorticoid excess. *Am. Heart J.* 86:404–14
107. Biglieri, E. G., Forsham, P. H. 1961. Studies on the expanded extracellular fluid and the responses to various stimuli in primary aldosteronism. *Am. J. Med.* 30:564–76
108. Chobanian, A. V., Burrows, B. A., Hollander, W. 1961. Body fluids and electrolyte composition in arterial hypertension: II. Studies in mineralocorticoid hypertension. *J. Clin. Invest.* 40:416–22
109. Tarazi, R. C., Dustan, H. P., Frohlich, E. D., Gifford, R. W. Jr., Hoffman, G. C. 1970. Plasma volume and chronic hypertension. *Arch. Intern. Med.* 125:835–42
110. Schalekamp, M. A. 1973. Hypertension in Chronic renal failure: An abnormal relation between sodium and the renin-angiotensin system. *Am. J. Med.* 55:379–90
111. Birkenhäger, W. H. et al 1972. Systemic and renal haemodynamics, body fluids and renin in benign essential hypertension with special reference to natural history. *Eur. J. Clin. Invest.* 2:115–22
112. Conn, J. W., Rovner, D. R., Cohn, E. L. 1968. Licorice-induced pseudo-aldosteronism. *J. Am. Med. Assoc.* 205:492–96
113. Koster, M., David, G. K. 1968. Reversible severe hypertension due to licorice ingestion. *N. Engl. J. Med.* 278:1381–83
114. Liddle, G. W., Bledsoe, R., Coppage, W. S. Jr. 1963. A familial renal disorder simulating primary aldosteronism but with negligible aldosterone secretion. *Trans. Assoc. Am. Physicians* 76:199–211
115. Christlieb, A. R., Gleason, R. E., Hickler, R. B., Lauler, D. P. 1974. Renin: A risk factor for cardiovascular disease? *Ann. Intern. Med.* 81:7–10
116. Stroobandt, R., Fagard, R., Amery, A.K.P.C. 1973. Are patients with essential hypertension and low renin protected against stroke and heart attack? *Am. Heart J.* 86:781–87
117. Gulati, S. C., Adlin, E. V., Biddle, C. M., Marks, A. D., Channick, B. J. 1973. The occurrence of vascular complications in low renin hypertension. *Ann. Intern. Med.* 78:828
118. Doyle, A. E., Jerum, G., Johnston, C. L., Louis, W. J. 1973. Plasma renin lev-

els and vascular complications in hypertension. *Brit. Med. J.* 2:206–7

119. Genest, J., Boucher, R., Kuchel, O., Nowaczynski, W. 1973. Renin in hypertension: How important as a risk factor? *Can. Med. Assoc. J.* 109:475–78

120. Veterans Administration cooperative study group on antihypertensive agents. 1967. Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. *J. Am. Med. Assoc.* 202:1028–34

121. Veterans Administration cooperative study group on antihypertensive agents. 1970. Effects of treatment on morbidity in hypertension II. Results in patients with diastolic blood pressure averaging 90 through 114 mm Hg. *J. Am. Med. Assoc.* 213:1143–52

Copyright 1975. All rights reserved

UNUSUAL ETIOLOGIES OF FEVER AND THEIR EVALUATION

❖7149

Sheldon M. Wolff, M.D., Anthony S. Fauci, M.D., and
David C. Dale, M.D.
Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20014

INTRODUCTION

Fever is one of the most common signs of illness, yet many febrile patients undergo initial medical evaluation without a diagnosis being established. If the fever lasts for at least 14 to 21 days in a patient who has been subjected to the usual office or outpatient examination and laboratory evaluation without a diagnosis being achieved, the provisional diagnosis of fever of unknown origin (FUO) is made. More extensive investigation of such patients will often lead to a diagnosis, although in many cases the fever disappears spontaneously before a diagnosis is made (1). A small group of patients with fever, however, defies extensive clinical and laboratory study and their fever persists for extended periods of time. During the past 14 years, we have had the opportunity to study over 200 patients with long-standing recurrent or chronic FUO, usually present for more than a year. The present chapter will outline the spectrum of our experience with this unusual population of patients.

GENERAL CONSIDERATIONS

Fever is a symptom of a vast array of illnesses. It is necessary to approach the evaluation of a patient with an FUO with patience, open-mindedness, and a willingness to pursue the diagnostic work-up as far as is feasible in accordance with sound clinical judgment. Of paramount importance is the documentation that the patient is, in fact, febrile. The work-up of such patients should be dictated by an organized and orderly approach. Evaluation should progress from the noninvasive (blood, urine, radiographic studies) to invasive procedures (bone marrow, liver biopsies, etc). Despite the reported diagnostic value of exploratory laparotomy in the usual FUO (2), it has been our experience that in the chronic FUO, which has been present for up to a year or longer, such an operative procedure is usually of minimal diagnostic value. With the availability of lymphangiography, peritoneoscopy, sophisticated endoscopies, and newer isotopic scanning techniques, a diagnosis previously made by laparotomy may now often be obtained by a less traumatic and safer procedure.

Fever is best defined as an increase in temperature over what is normal for a given individual at that particular time of day and not merely an isolated temperature greater than 37° C (98.6° F). In careful monitoring of circadian temperature rhythms of normal young adults on our wards over a three-month period, the mean oral temperature was 97.8° F, with a mean increase from morning to evening of 0.8° F.

INFECTIOUS DISEASES

Infectious agents are the most common causes of FUO, but chronic fevers are less likely to be of microbial origin. Nonetheless, infection should be a prime consideration in all febrile patients regardless of duration of fever. In our patients some of the infectious etiologies of recurrent, long-standing fevers have been such unexpected conditions as: (*a*) osteomyelitis of the leg, which had been considered by the physician to be sciatic nerve root compression. (*b*) subphrenic abscess two years after removal of a hydatid cyst of the liver; (*c*) *Plasmodium ovale* malaria in a patient who had left an endemic area three years before, had never had a recognized attack, had regularly taken chemoprophylaxis while in the endemic area, and had received adrenal corticosteroids for a tendonitis two weeks before the malaria became clinically apparent; (*d*) trypanosomiasis in an Ugandan forester who had been having episodic fevers for four and one half years without any central nervous system involvement; and (*e*) Whipple's disease in a patient with four years of recurrent fever and a diagnosis obtained through intestinal biopsy. Obviously, to make such diagnoses, one must first be prepared to consider such unlikely conditions. Apart from these unusual instances of infections mentioned above, the general rule has clearly been that as the duration of fever increases, the likelihood of an infectious cause decreases.

NEOPLASTIC DISEASES

Like infectious diseases, a neoplastic process is less likely as the cause of FUO the longer the fever has been present. In our experience, lymphomatous disease rarely is the cause of an FUO of longer than one year's duration, but one should always consider this possibility since we have had a patient with a lymphoma who had a two and one half year illness characterized by fever and splenomegaly before the diagnosis became apparent. Splenic scan demonstrated an intrasplenic defect and splenectomy was performed for a presumptive splenic abscess, which turned out to be a necrotic lymphosarcoma. Adenocarcinomas are not one of the neoplastic diseases considered to be associated with fever. However, we routinely perform barium enemas on our long-term FUOs since we have had two patients with one and one half and three years of recurrent fever, respectively, who had adenocarcinomas of the colon. Removal of the primary lesion led to cessation of the fever in both cases.

"EXAGGERATED" CIRCADIAN TEMPERATURE RHYTHM

A large number of young children of both sexes, as well as several young female adolescents, have been evaluated because of FUO. These individuals have often had a previous acute, self-limited febrile illness of infectious origin. The patient or the family then becomes involved in the frequent monitoring of temperature. Following recovery from this illness, it is noted that the temperature (usually in the early evening) never returns to "normal." These individuals seem to have exaggerated daily swings and their normal evening temperature may be 99–100° F. After a minimal work-up, which is completely normal, we observe such individuals and encourage them to not take their temperature unless they are ill. Unfortunately, many of these individuals have been subjected to unnecessarily expensive, painful, and often dangerous procedures when, in fact, they are normal. Many years of follow-up have substantiated our opinion that these patients were not ill, but merely had this exaggerated daily swing in temperature brought to attention by a routine, self-limited infectious illness.

FACTITIOUS ILLNESS

The more patients one sees with chronic, unexplained fevers, the more likely is one to encounter patients with factitious illnesses. Almost 10% of patients referred to us with chronic FUO have had such diseases. Aside from a few young males, none older than adolescence, all of these patients were young, adult females and with few exceptions were in the allied health professions, such as nursing. A few were injecting themselves with materials such as milk or bacteria, causing a true inflammatory or microbial-related fever. However, most patients were manipulating or changing thermometers and, thus, simulating a true fever. The most important factor in recognizing these patients is an awareness of the condition. We have found that most patients perform such acts as a means of asking for help and that, in fact, the majority are relieved when their deception is discovered. The recognition and confrontation of such patients requires timing, patience, tact, and understanding. Once the patient is confronted, psychiatric therapy usually proves to be very beneficial (3).

GRANULOMATOUS DISEASES

The largest numbers of FUO patients whom we have seen fall into the category of granulomatous diseases. For example, we have studied seven patients with recurrent, high spiking fevers, some without gastrointestinal symptoms, who clearly had regional enteritis documented by small bowel series.

Close to 10% of all patients admitted to our service with recurrent, long-standing fever have had granulomatous hepatitis as an isolated disease entity (4). These patients were usually in the fifth to sixth decade of life, had an average of three and one half years of unexplained fever, and only mild to moderately abnormal liver function tests. In fact, many of these patients had normal liver function tests, despite the presence of granulomata on liver biopsy. Extensive evaluations provided no evidence to support a diagnosis of sarcoidosis. Therefore, if initial work-up of prolonged FUO fails to provide a diagnosis, we perform percutaneous liver biopsies in our patients. It is important to recognize this syndrome of granulomatous hepatitis, since these patients usually respond to anti-inflammatory therapy. Other granulomatous diseases which may be seen associated with prolonged fevers include Takayasu's arteritis and temporal arteritis. The latter diagnosis can be made by biopsy of the temporal artery. Any patient with hectic fevers, severe myalgias, anemia, and very high erythrocyte sedimentation rate should be suspected of having this diagnosis (5).

RHEUMATIC DISEASES

Despite careful evaluation with numerous serological tests in all of our patients, the diagnosis of systemic lupus erythematosus was rarely made. However, we have seen long-term fevers as a manifestation of hypersensitivity vasculitides, polyarteritis nodosa, and erythema multiforme. A common condition which we have seen presenting with fever of unknown origin, arthralgias, sore throat, serositis, skin rash, lymphadenopathy, and hepatosplenomegaly, is juvenile rheumatoid arthritis (Still's disease). This diagnosis is frequently entertained and relatively easy to make in children. However, we have now seen 14 adult patients with an identical syndrome as that seen in children (6). It is important that this syndrome be recognized in adults since salicylate therapy usually results in dramatic clinical remissions and control of disease activity.

METABOLIC AND INHERITED DISEASES

We have diagnosed unusual metabolic and inherited diseases in a small number of patients referred for evaluation of long-standing FUO. For example, a patient with recurrent episodes of fever, abdominal pain, and hepatosplenomegaly was found to have hypertriglyceridemia and hypercholesterolemia. Reduction of his saturated fat intake resulted in disappearance of the fever and symptoms. Another metabolic disorder which has presented as an FUO is Fabry's disease. Careful evaluation of this patient revealed angiomatous skin lesions, and biochemical studies of the urine established this diagnosis. Familial Mediterranean fever (FMF) is easily recognizable in some ethnic groups, yet it is frequently forgotten that it can occur in other groups as well. At least six patients of other extractions such as German, Irish, and French, have been referred to us with recurrent episodes of fever, peritonitis, and leukocytosis, consistent with classical FMF. If these patients had been Arabs, nonAshkenazic Jews, or Armenians, there would have been no difficulty in making a diagnosis of FMF. Yet, misconceptions regarding the restricted ethnic distribution led to delay in making the correct diagnosis.

Of interest is the fact that of all the patients we have seen, the only patients with a true periodic or cyclical illness are those with cyclic neutropenia (7). These patients have their fever at 21-day intervals at the time of maximal neutropenia. The diagnosis can be established only by doing serial white blood cell differential counts.

Despite the fact that it has been shown that there is no relationship between plasma etiocholanolone levels and fever of unknown origin (8), the diagnosis of etiocholanolone fever is still considered by some physicians. There is no solid evidence at present to support such a diagnosis in any patient.

THERMOREGULATORY DISORDERS

Although several patients have been referred to us with a proposed diagnosis of "central" or "hypothalamic" disorders, these disorders are extremely rare. We have seen a few patients who have had long-standing, high fevers initiated by a seemingly mild upper respiratory tract infection. Studies during artificial environmental cold stress in these patients showed that they did not maintain their normal body temperatures and reacted with a decrease in temperature similar to that anticipated in a poikilothermic animal. In addition, some patients have had recurring episodes of fever, endocrine abnormalities, and thermoregulatory disturbances. Such patients, in particular one young male (9), have had very gratifying therapeutic responses to chlorpromazine. It is to be emphasized that the latter drug should be reserved only for those patients in whom such a rare diagnosis seems clear.

CONCLUSIONS

Study of a large group of patients with recurrent or prolonged fevers of unknown etiologies has resulted in a wide spectrum of diagnoses. In many patients the diseases themselves are rare, while in a few patients the prolonged fever is an unusual manifestation of a relatively common disease. Such patients provide a real challenge to physicians, and the results of meticulous and orderly work-ups in these patients can often lead to resolution of difficult diagnostic problems and, in many, to successful therapy.

Litertaure Cited

1. Petersdorf, R. G., Beeson, P. B. 1961. Fever of unexplained origin. Report on 100 cases. *Medicine* 40:1–30
2. Geraci, J. E., Weed, L. A., Nichols, D. R. 1959. Fever of unexplained origin. The value of abdominal exploration in diagnosis. *J. Am. Med. Assoc.* 169:1306–15
3. Herzberg, J. H., Wolff, S. M. 1972. Chronic factitious fever in puberty and adolescence. A diagnostic challenge to the family physician. *Psychosom. Med.* 3:205–12
4. Simon, H. B., Wolff, S. M. 1973. Granulomatous hepatitis and prolonged fever of unknown origin: A study of 13 patients. *Medicine* 52:1–21
5. Wilske, K. R., Healy, L. A. 1967. Polymyalgia rheumatica. A manifestation of systemic giant-cell arteritis. *Ann. Intern. Med.* 66:77–86
6. Bujak, J. S., Aptekar, R. G., Decker, J. L., Wolff, S. M. 1973. Juvenile rheumatoid arthritis presenting in the adult as fever of unknown origin. *Medicine* 52:431–44
7. Guerry, D., Dale, D. C., Omine, M., Perry, S., Wolff, S. M. 1973. Periodic hematopoiesis in human cyclic neutropenia. *J. Clin. Invest.* 52:3220–30
8. George, J. M., Wolff, S. M., Diller, E., Bartter, F. C. 1969. Recurrent fever of unknown etiology. Failure to demonstrate association between fever and plasma unconjugated etiocholanolone. *J. Clin. Invest.* 48:558–63
9. Wolff, S. M., Adler, R. C., Buskirk, E. R., Thompson, R. H. A syndrome of periodic hypothalamic discharge. *Am. J. Med.* 36:956–67

Copyright 1975. All rights reserved

CELLULAR KINETICS, STAGING, AND IMMUNOGLOBULIN SYNTHESIS IN MULTIPLE MYELOMA[a]

❖7150

Brian G. M. Durie, M.D. and Sydney E. Salmon, M.D.
Section of Hematology and Oncology, Department of Internal Medicine, University of Arizona, College of Medicine, Tucson, Arizona 85724

Multiple myeloma has served as a model of neoplastic disease in both mouse and man (1). Investigations of mouse myeloma first demonstrated the quantitative relationship between the concentration of serum myeloma proteins and tumor weight (1). This was the basis for calculation of myeloma cell mass in patients with IgG myeloma from measurements of M-component synthesis and metabolism (2). Serial determinations of M-component concentration served to provide cell mass values for the kinetic analysis of tumor growth and regression with treatment (3). From experiments with various animal tumors it has become apparent that such information is critical to the design of optimal chemotherapy schedules (4–6). Since it is rarely possible to make direct serial measurements of total tumor mass in man, use of the "tumor marker" approach provides new and important information about tumor growth and regression in individual patients. In this review we shall describe the approach which we have developed to determine the cellular kinetics of multiple myeloma, as well as to stage myeloma patients and assess the response to chemotherapy, all based primarily upon quantitative studies of immunoglobulin M-component synthesis by human myeloma cells in vitro and in vivo.

M-COMPONENT SYNTHESIS AND MYELOMA TUMOR CELL NUMBER

Initial studies were carried out to determine the average cellular synthetic rate of M-component by freshly aspirated myeloma cells in vitro. Our radioimmunoassay and tissue culture techniques for invesitgation of patients with serum M-components

---

[a]Supported by U.S. Public Health Service grant CA-14102 and by contract N01-CM-33713 from the National Cancer Institute, National Institutes of Health.

have been reported in detail (2, 7, 8) and recently reviewed (9). Over 120 myeloma patients have been studied in the past five years, including many serial studies, comprising in excess of 400 synthetic rate experiments. All myeloma types with the exception of IgD have had the in vitro immunosynthetic tests performed. We have found that the average cellular M-component synthetic rate may be as low as 2.5 pg per myeloma cell per day, or as high as 38 pg per cell per day, with most patients being in the 10–20 pg per day range. Expressed in molecular terms, rates ranged from 5200 to 90,000 molecules of immunoglobulin secreted per cell per minute, while the intracellular content of M-component ranged from 5–10 million molecules per cell. Thus, these cells can have enormous immunosynthetic capacity, and secrete an amount of immunoglobulin equal to their intracellular content within 2–4 hr. Myeloma cases with clinical evidence of hyperviscosity were definitely in the high immunosynthetic rate range (24–28 pg/cell/day).

The total body production rates for M-components were initially measured in vivo using metabolic turnover techniques with labeled immunoglobulins (2). Because of the availability of extensive studies of this nature (10), it is now possible to calculate the total body production rate with reasonable accuracy from the serum M-component concentration, the body weight, and the hemoglobin. When available, measurements of the plasma volume provide further accuracy. In patients with very high serum M-component concentrations or hyperviscosity, serial plasma volumes are particularly important.

A simple mathematical relationship was developed to relate these M-component synthetic rates and derive the total body myeloma cell number:

$$\text{Total Body Myeloma Cell Number} = \frac{\text{Total Body M-Component Synthetic Rate.}}{\text{Cellular M-Component Synthetic Rate}}$$

Total body myeloma cell numbers were found to be in the range of 0.2×10^{12} cells to 3×10^{12} cells. There was a clear quantitative relationship between the myeloma cell number and the clinical extent of disease, such as severity of lytic bone involvement. Although it was only rarely possible to confirm this calculated tumor burden in man, Hiramoto et al (11), using a subcutaneous plasmacytoma model in mice, have confirmed the general validity of the approach by showing that tumor burden and drug-induced changes in tumor mass as calculated from serum M-component levels can be very closely correlated with the measured weights of the subcutaneous plasmacytomata. These observations not only extend earlier work (1, 12), but also validate the use of this approach in clinical oncology. From the studies in patients with IgG myeloma (2), it was noted that clinical diagnosis is never "early" because at least 0.2×10^{12} (about 200 g) myeloma cells are present in the body in "early" cases. In most instances, well over 1×10^{12} myeloma cells are present at the time of diagnosis, and patients with multiple lytic bone lesions have in excess of 2×10^{12} tumor cells in the body. A lethal burden of myeloma cells is in the range of 5–7% of body weight (assuming 10^{12} myeloma cells = 1 kg).

CELLULAR KINETICS OF GROWTH AND REGRESSION

From serial measurements of serum M-component and other parameters of M-component metabolism it was possible to construct a plot of calculated myeloma cell mass with time. Assuming constant exponential growth, it was evident (by extrapolation) that the preclinical course of myeloma could extend over as much as 15–20 yr. This was similar to suggestions previously made by Hobbs (13). However, it was not at all clear that assumption of a constant growth rate was reasonable. Virtually all animal tumors can be better described with a Gompertzian growth curve (14, 15), in which growth is initially exponential, but slows progressively as the tumor enlarges. More detailed analyses of growth curves in untreated patients and the regression curves after chemotherapy were therefore carried out. Use of a computer system was essential in this work. The results indicated that the growth mode of myeloma was Gompertzian. Although the tumor had doubling times of 4–6 months at the time of clinical presentation, by back calculation from a patient's data it could be discerned that the initial doublings of the clone took from 1–3 days. The short initial doubling time is consistent with the measured generation time of myeloma cells, which is generally less than 72 hr (16, 17). We now believe that myeloma growth occurs over 1–2 yr and has almost plateaued at the time of clinical presentation. Indeed, some benign monoclonal gammopathies may represent monoclones that have plateaued at a sublethal level.

It was also of great interest to find that initial exponential regression followed by a plateaued response to chemotherapy could be fitted to the Gompertz equation (3). This characteristic mode of regression was noted after both continuous and pulse (every 4–6 weeks) melphalan therapy. The tumor regresses rapidly initially, and then slows and reaches a plateau size which is stable for a period of months or years even though treatment is continued. This implies that the fraction of proliferating myeloma cells (those undergoing DNA synthesis) increased as the tumor mass became smaller (3). Subsequently, we tested that possibility by measuring the tritiated thymidine labeling index of myeloma cells before and after melphalan therapy (18, 19). We found that the proliferating component (growth fraction) of the tumor did increase as the tumor regressed and plateaued in size, suggesting that one or more of the cycle-active drugs (e.g. cytosine arabinoside, methotrexate, azathioprine, hydroxyurea, vincristine) might have a role in myeloma therapy after the first few logs of non-proliferating myeloma cells had been eliminated. This hypothesis is now undergoing testing in the course of current clinical studies.

DEVELOPMENT OF A CLINICAL STAGING SYSTEM

It was thought that this type of kinetic analysis could be of major importance in myeloma treatment protocol studies, both for developing study designs and for interpreting results. To enable widespread use of the system, a method of clinical staging was essential for institutions unable to carry out the detailed in vitro studies. Our patients who had had such studies were analyzed in detail and a clinical staging system in which the presenting clinical features were correlated with measured myeloma cell burden was developed (20). This system is illustrated in Table 1. By back calculation it was possible to derive the in vitro M-component synthetic rate. This was then used as the basis for serial follow-up.

The whole system (Myeloma Study System—MSS) was set up as part of a General Electric time-sharing network available to any user in the USA or thirteen foreign countries (21). Using a conversational mode, all basic clinical details of a myeloma patient could be entered into the system. The computer routines automatically

Table 1 Clinical staging system for multiple myeloma

Stage	Criteria	Myeloma Cell Mass (cells × $10^{12}/M^2$)[a]
I	ALL of: 1. Hemoglobin > 10 g/100 ml 2. Serum calcium normal 3. On X ray, normal bone structure or solitary lytic lesion only 4. Low M-component production rates a. IgG < 5 g/100 ml b. IgA < 3 g/100 ml c. Urine light chain < 4 g/24 hr (electrophoresis)	< 0.6×10^{12} (Low)
II	Fitting neither I or III	0.6 - 1.2×10^{12} (Intermediate)
III	ONE or MORE of: 1. Hemoglobin < 8.5 g/100 ml 2. Serum calcium > 12 mg/100 ml 3. Advanced lytic bone lesions 4. High M-component production rates a. IgG > 7 g/100 ml b. IgA > 5 g/100 ml c. Urine light chain > 12 g/24 hr (electrophoresis)	> 1.2×10^{12} (High)
A or B	A = Relatively normal renal function (BUN ≤ 30 mg/100 ml, creatinine ≤ 2 mg/100 ml) B = Clearly abnormal renal function (BUN > 30 mg/100 ml, creatinine > 2 mg/100 ml)	
	Examples: A patient staged as IB has a low tumor cell mass and impaired renal function. A patient staged as IIIA has a high tumor cell mass and normal renal function.	

[a] 10^{12} cells — approximately 1 kg or 2.2 lb; M^2 = square meter of body surface area.

calculated the myeloma cell mass and estimated M-component synthetic rate with instant feedback to the user. The system was designed so that all interaction could be accomplished by a physician or medical secretary. With serial follow-up entries it is now possible to plot the change in myeloma cell mass with time. The percentage tumor regression and whether response has been achieved as per specified criteria [e.g. Chronic Leukemia-Myeloma Task Force Guidelines (22)] can be printed out. Comparisons of treatment groups, drug schedules, clinical and laboratory "risk factors," or any of a number of parameters can be achieved without difficulty. Thus, the system sets the stage for predictive cancer chemotherapy in individual patients who are closely followed. Consequently, protocols of the future may *not* be uniform for different patients; they could be individualized with a computer-aided decision for timing, sequence, and combination of drugs, to be used to maximally exploit the kinetic sensitivity of the tumor to drugs.

Used in this fashion, the serum M-component is an excellent and helpful marker of myeloma cell mass. However, there are other markers of tumor mass which may prove valuable in a large variety of other tumors. Some of these are listed in Table 2. For example, although carcinoembryonic antigen (CEA) has failed to be as useful as was initially thought might be possible in early diagnosis, preliminary work does suggest that CEA may be useful as a marker of extent of disease and prognosis (23, 24). Also, polyamines, which are present in increased levels in serum and urine in patients with a wide variety of neoplasms, may well prove to be very useful markers, both of extent of tumor mass as well as degree of cell kill with therapy (25, 26).

The "marker kinetic" approach as applied in multiple myeloma thus serves as a model for the more widespread application of the technique. It provides a direct means of testing the value of cellular kinetics for planning and evaluation of cancer therapy.

Table 2 Examples of tumor markers

Marker	Tumor
Carcinoembryonic antigen (CEA)	Gastrointestinal, lung, breast
Alpha-feto protein	Hepatoma, testicular, ovarian, stomach
Placental alkaline phosphatase (Regan)	Lung
Muramidase	Leukemia
Acid phosphatase	Prostate
Human chorionic gonadotrophin (HCG)	Trophoblastic
Adrenocorticotrophic hormone (ACTH)	Lung, thymus, pancreas
Erythropoietin	Renal
Parathormone	Lung, kidney, lymphoma
Antidiuretic hormone (ADH)	Lung
5-Hydroxyindoleacetic acid (5HIAA)	Carcinoid
Catecholamines	Pheochromocytoma, neuroblastoma
Polyamines	Many (e.g. lymphomas, leukemias, colon, breast)

Literature Cited

1. Nathans, D., Fahey, J. L., Potter, M. 1958. The formation of myeloma protein by a mouse plasma cell tumor. *J. Exp. Med.* 108:121
2. Salmon, S. E., Smith, B. A. 1970. Immunoglobulin synthesis and total body tumor cell number in IgG multiple myeloma. *J. Clin. Invest.* 49:1114
3. Sullivan, P. W., Salmon, S. E. 1972. Kinetics of tumor growth and regression in IgG multiple myeloma. *J. Clin. Invest.* 51:1697
4. Skipper, H. E. et al 1970. Implications of biochemical, cytokinetic, pharmacologic, and toxicologic relationships in the design of optimal therapeutic schedules. *Cancer Chemother. Rep.* 54:431
5. Schabel, F. M. Jr. 1969. The use of tumor growth kinetics in planning "curative" chemotherapy of advanced solid tumors. *Cancer Res.* 29:2384
6. Goldin, A. 1969. Factors pertaining to complete drug-induced remission of tumor in animals and man. *Cancer Res.* 29:2285
7. Salmon, S. E., Mackey, G., Fudenberg, H. H. 1969. "Sandwich" solid phase radioimmunoassay for the quantitative determination of human immunoglobulins. *J. Immunol.* 103:129
8. Salmon, S. E., Smith, B. A. 1970. Sandwich solid phase radioimmunoassays for the characterization of human immunoglobulins synthesized in vitro. *J. Immunol.* 104:665
9. Salmon, S. E. 1973. Immunoglobulin synthesis and tumor kinetics of multiple myeloma. *Semin. Hematol.* 10:135
10. Waldmann, T. A., Strober, W. 1969. Metabolism of immunoglobulins. *Progr. Allergy* 13:1
11. Ghanta, V. K., Hiramoto, R. N. 1974. Quantitation of total-body tumor cells (MOPC 104E). I. Subcutaneous tumor model. *J. Nat. Cancer Inst.* 52:1119
12. Hobbs, J. R. 1971. Immunocytoma o' mice an' men. *Brit. Med. J.* 2:67
13. Hobbs, J. R. 1969. Growth rates and responses to treatment in human myelomatosis. *Brit. J. Haematol.* 16:607
14. Laird, A. K. 1964. Dynamics of tumor growth. *Brit. J. Cancer* 18:490
15. Laird, A. K. 1965. Dynamics of tumour growth: Comparison of growth rates and extrapolation of growth curve to one cell. *Brit. J. Cancer* 19:278
16. Killmann, S. A., Cronkite, E. P., Fliedner, T. M., Bond, V. P. 1962. Cell proliferation in multiple myeloma studied with tritiated thymidine in vivo. *Lab. Invest.* 11:845
17. Queisser, W., Hoelzer, D., Queisser, U. 1973. Cell proliferation in leukaemic forms of paraproteinaemic haemoblastosis studied by combined cytophotometric and autoradiographic technique. *Klin. Wochenschr.* 51:230
18. Salmon, S. E., Smith, B. A. 1972. Induction of tumor-susceptibility to cycle-active agents in IgG multiple myeloma. *Clin. Res.* 20:570
19. Salmon, S. E. 1975. Expansion of the growth fraction in multiple myeloma with alkylating agents. *Blood* 45:1
20. Salmon, S. E., Durie, B. G. M. 1975. Cellular kinetics in multiple myeloma. A new approach to staging and treatment. *Arch. Intern. Med.* 135:1
21. Durie, B. G. M., Lee, B. J., Salmon, S. E. 1974. A new simple staging system for multiple myeloma. *Proc. Am. Assoc. Cancer Res.* 15:90 (#358)
22. Chronic Leukemia-Myeloma Task Force, N.C.I. Proposed guidelines for protocol studies. II. Plasma cell myeloma. 1973. *Cancer Chemother. Rep.* Part 3, Vol. 4:145
23. Holyoke, D. G., Reynoso, T., Chu, T., Murphy, G. 1972. Carcinoembryonic antigen in gastrointestinal carcinoma. *Clin. Res.* 20:567 (Abstr.)
24. Dhar, P., Moore, T., Zamchek, N., Kupchik, H. Z. 1972. Carcinoembryonic antigen in colonic cancer: use in pre-operative and post-operative diagnosis and prognosis. *J. Am. Med. Assoc.* 221:31
25. Russell, D. H., Ed. 1973. Polyamines in Normal and Neoplastic Growth. New York: Raven
26. Russell, D. H., Gullino, P., Marton, L. J., LeGendre, S. 1974. Spermidine elevations in sera of rats with regressing mammary tumors. *Proc. Am. Assoc. Cancer Res.* 15:5 (#19)

Copyright 1975. All rights reserved

GRANULOCYTE TRANSFUSION THERAPY

❖7151

Donald J. Higby, M.D. and Edward S. Henderson, M.D.
Department of Medicine A, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, New York 14203

INTRODUCTION

The potential of chemotherapy and radiotherapy to effect long term remissions and even cures of disseminated malignancies has greatly increased over the past few years. Patients receiving nonsurgical cancer therapy almost invariably develop bone marrow depression, often to a clinically significant degree. Although red cell and platelet transfusions are effective and readily available to alleviate anemia and thrombocytopenic hemorrhage, prevention and control of infection related to neutropenia remains a major clinical problem.

In the recent experience of the National Cancer Institute, infection was the main or complicit cause of death in 79% of 450 patients with hematologic malignancies or aplastic anemia (1). A recent review of causes of death in lymphoma showed that 51% of 164 patients died of infection (2). Fifteen to forty percent of patients with metastatic carcinoma die with identifiable infection (3). In addition, the management of malignant disease is often complicated by nonfatal infection, which imposes limitations on the application of effective chemotherapeutic regimens.

Various approaches have been used in an attempt to control infection. Multidrug antibiotic regimens have been advocated (4–6) but there is as yet no clear-cut evidence that such approaches are in themselves sufficient. Several reports on the use of "germ free" isolator systems with or without gut sterilization during remission induction of acute leukemia suggest that their use results in delay in the development of infection and an overall reduction in frequency of severe infection (7–9). However, these measures are costly and cumbersome and are of no value in the treatment of established infection.

Heuristically, the replacement of white cells in neutropenic patients would seem to be of value since there is a clear correlation between the degree of granulocytopenia and acquisition of infection (10). To date, however, this form of therapy has not been generally available and many therapists consider the procedure entirely experimental. Nevertheless, the control of infection in neutropenic patients will almost certainly be related to our abilities to obtain and transfuse large quantities of normal granulocytes.

HISTORY

Although Strumia is often given credit for the earliest attempts at granulocyte transfusions, he in fact injected leukocytes intramuscularly in the hope of stimulating the production of new leukocytes in the marrow rather than as replacement therapy (11). Brecher and co-workers (12) demonstrated that granulocytes infused into irradiated aplastic dogs migrated to areas of infection, suggesting the feasibility of such transfusions on a large scale in humans. In 1961, Freireich's group showed that transfusion of cells obtained from donors with chronic granulocytic leukemia (CGL) resulted in defervescence and resolution of infection in granulocytopenic recipients (13). Yankee et al in 1964 obtained large numbers of normal granulocytes by combining buffy coats from 50 to 100 units of whole blood obtained from 15 to 25 donors (14). Recovery and clinical response in recipients resembled what had been noted in patients receiving CGL cells.

In 1965, Freireich, Judson & Levine (15), representing a joint venture of the National Cancer Institute and the International Business Machine Corporation, described the development of a continuous flow centrifugation (CFC) system which allowed for more convenient leukocyte separation. However, granulocyte yields were initially low and red cell contamination was high, so that resultant leukocyte doses were small. In 1969, Buckner applied the technique to volunteer donors with CGL and was able to obtain between 1.2 and 19.4 X 10^{11} CGL leukocytes per donation (16). McCredie reported in 1971 that granulocyte yields from normal donors could be increased 10—fold through the use of etiocholanolone to increase peripheral granulocyte numbers and hydroxyethyl starch to increase sedimentation of red blood cells (17).

Finally, Djerassi (18) described a method of reversible leukoadhesion ("filtration" leukapheresis—FL) which allows for practicable and inexpensive collection of large quantities of granulocytes from normal donors and which can be adapted to almost any blood bank.

KINETICS OF THE MYELOID COMPARTMENT

Normal Kinetics

Rational granulocyte replacement therapy requires an understanding of neutrophil kinetics. In the marrow, stem cells give rise to committed myeloid progenitors which are mitotically active for two or three days (myeloblasts, promyelocytes, myelocytes). They then enter the postmitotic compartment (metamyelocytes, bands, neutrophils), still in the marrow. A maturation process of about ten days takes place before mature cells are discharged into the peripheral blood. About half the granulocytes marginate on small blood vessel walls. Neutrophils in the peripheral circulation enter the tissues in a random fashion. Under severe stress of infection, transition through the postmitotic compartment can be reduced to as little as 48 hr (19).

Granulocytes leave the circulation with a half time of 6.7 hr in normal circumstances (20).

The normal adult requires an input of 10^{11} polymorphonuclear neutrophils (PMN) per day to maintain a steady state. This is equivalent to the number of PMN contained in 6–8 blood volumes (21). The size of various compartments is shown in Table 1.

Table 1 The average size of various granulocyte compartments[a]

Total blood granulocyte pool	7×10^8 cells/kg
Marginated pool	3.1×10^8 cells/kg
Circulating pool	3.9×10^8 cells/kg

[a] Modified from Cartwright, G.E. et al 1964. *Blood* 24:784.

Pharmacologic Alterations in the Myeloid Compartment

Many drugs affect the granulocyte pool through idiosyncratic or immunologically mediated reactions. In addition, some agents used for purposes other than cancer chemotherapy are directly toxic to the marrow at high doses. The most toxic agents in clinical use, however, are the drugs developed for cancer chemotherapy. Table 2 shows the relative toxicity of commonly used drugs to the granulocyte pool.

A single dose of drug capable of damaging only dividing cells, and thus unable to injure most pluripotential stem cells or cells in the postmitotic compartment, will result in a nadir of neutropenia about 8 to 10 days following administration, with rapid recovery as the myeloid compartment is repopulated with the progeny of uninjured stem cells. Drugs capable of injuring nondividing cells as well as dividing cells cause delay in recovery from granulocytopenia proportionate to the degree of injury to the stem cell compartment. The nitrosoureas are a class of compounds in which injury is primarily to the stem cell compartment. The reflection of this injury by granulocytopenia occurs four to six weeks after administration of a single dose of the drug (19).

Table 2 Myelosuppressive potential of commonly used chemotherapeutic agents[a]

Rarely cause granulocytopenia:
  steroids, androgens and estrogens, asparaginase
Myelosuppression not dose-limiting:
  vincristine, bleomycin
Myelosuppression, a dose-limiting factor, but other toxicity equally important:
  procarbazine, mithramycin, cyclophosphamide, 5-fluorouracil, methotrexate, actinomycin D, 6-mercaptopurine
Primary myelosuppressive toxicity:
  cytosine arabinoside, alkylating agents, vinblastine, nitrosoureas

[a] Henderson, E. S. In *Drugs and Hematologic Reactions,* ed. N. V. Dimitrov, J. H. Nodine, 208. New York: Grune & Stratton.

Pathologic Alterations in the Myeloid Compartment

INFECTION In the early stage of infection, granulocytes leave the marrow and traverse the blood compartment more rapidly than usual, causing an initial increase in the marginated pool followed by an increase in the circulating pool. More rapid entry into the tissue occurs. There initial events are most likely due to humoral factors. About 13 hr later, the total blood granulocyte pool is at a peak while the transit time has returned to normal. During recovery of infection, transit time is prolonged (22, 23). Infection in individuals with bone marrow compromise may result in neutropenia and depletion of marrow stores (24).

HEMORRHAGE The same stem cell population gives rise to erythrocytes, granulocytes, and megakaryocytes. In animals made plethoric by hypertransfusion, an increased number of PMN and their precursors are evident. Conversely, animals subject to bleeding are noted to have a reduced number of PMN and their precursors (23).

ALLERGIC REACTIONS The half time of transfused labeled autologous granulocytes is longer than that of labeled ABO and Rh compatible isologous granulocytes in human subjects not previously exposed to blood products. This may be related to naturally occurring isoagglutinins (25). When leukocyte-rich blood fractions are transfused into patients with histories of febrile transfusion reactions (26), or when HL-A incompatible platelets are transfused into sensitized recipients (27), a prolonged decrease in circulating granulocytes is noted. The source of replacement granulocytes during recovery is the bone marrow. Adsorption of antigen-antibody complexes by recipient PMN and subsequent clearance of these PMN by the reticuloendothelial system probably accounts for these observations (27).

Mechanisms for Increasing the Circulating Granulocyte Pool

The circulating granulocyte pool can be increased by: (*a*) inhibiting the entry of circulating PMN into the tissue compartment, (*b*) increasing the rate of release from marrow, or (*c*) increasing the circulating pool at the expense of the marginated pool.

EPINEPHRINE Administration of epinephrine results in an abrupt elevation of circulating PMN, probably through demargination. This is followed by neutropenia as circulating cells are sequestered in tissues. A more sustained elevation occurs beginning 2 to 4 hr after the drug is given and lasts 3 to 6 hr as new cells are released from marrow (28).

GLUCOCORTICOID HORMONES The observed increase in circulating granulocyte count following the administration of glucocorticoid hormones (29, 30) is partly due to fluid shifts and relative lymphopenia. In addition, release of PMN from the marrow compartment is increased (31) and egress from the vascular compartment decreased (32). Peak granulocyte levels occur 4 to 6 hr after an oral dose. Granulocytes from a donor who has received glucocorticoids kill bacteria in an in vitro test system as well as those from an untreated donor (33).

ETIOCHOLANOLONE Etiocholanolone, a naturally occurring metabolite of dehydroepiandrosterone, produces fever and leukocytosis in man. This leukocytosis is not at the expense of the marginated pool, and it appears to result solely from mobilization of the marrow reserve (34). The peak increase in granulocyte count occurs about 8 to 14 hr after administration of the drug (35).

ENDOTOXIN Endotoxin injected intravenously produces an initial decrease in the number of circulating neutrophils followed by a rebound granulocytosis. An endotoxin-complement product most likely is ingested into neutrophils resulting in cellular injury associated with decreased metabolic activity and leakage of kinase and lysozyme (36). The injured PMN are cleared by the reticuloendothelial system. Products of neutrophil damage act either directly or mediately to cause release of bone marrow granulocyte reserves (37). As with etiocholanolone, peak neutrophilia is delayed.

OTHER DRUGS Other drugs capable of causing transient neutropenia followed by neutrophilia include histamine, nicotinic acid, and dextran (38).

PHYSICAL METHODS If blood from normal volunteers is incubated in a hemodialysis coil and reinfused, immediate profound and very brief neutropenia occurs. This is followed by a sustained increase in the circulating granulocyte count (39). Simple phlebotomy and reinfusion does not duplicate the phenomenon. The neutropenia is due to pulmonary bed sequestration (40) while the PMN elevation is due to bone marrow release (38). Similar observations have been noted in donors undergoing FL (41) or receiving infusions of disrupted leukocytes (42). The sequence of events leading to these changes most likely begins with alterations in PMN membranes leading to their pulmonary sequestration and release of humoral factors, which in turn cause increased discharge of granulocytes from the marrow.

Physical and chemical manipulation of the granulocyte pool in donors can be used to increase the efficiency of cell harvesting procedures, all of which are ultimately dependent on the number of circulating PMN. Some methods in current clinical use will be described later.

THE EFFICACY OF GRANULOCYTE TRANSFUSION IN ANIMAL EXPERIMENTS

In total-body-irradiated dogs innoculated with *Escherichia coli*, single homologous leukocyte transfusions were shown to prolong significantly the duration of blood sterility and subsequent survival (43). Similar experiments confirmed these observations in animals made neutropenic by chemotherapy and allowed to acquire infections naturally (44). Nontransfused dogs developed septicemia and died within seven days. Three of six transfused dogs survived while three died of thrombocytopenic hemorrhage. The latter had negative blood cultures throughout. Dale (45) developed a model for the study of granulocyte transfusions in dogs made neutropenic by total body radiation. Pneumonia was induced by injecting 5×10^8 *Pseudomonas* bacteria into one main stem bronchus during light anesthesia. No apparent difference in survival was noted in dogs receiving antibiotics vs dogs in the control group. Eight pairs of dogs were treated in which one of each pair received gentamicin and the other received gentamicin plus white cell transfusions. Significantly improved survival was noted in the transfused group, which contained the only two long term survivors, both of whom subsequently recovered full hematopoietic function.

THE EFFICACY OF GRANULOCYTE TRANSFUSIONS IN HUMANS

Clinical results of granulocyte transfusion therapy have been reported by many investigators. Most reports suffer from the lack of a distinct end point, and it is difficult to draw conclusions concerning the role of granulocytes in patients who are experiencing alterations in bone marrow status during the period of observation, and who are also receiving a variety of antibiotics, antipyretics, and steroids. Furthermore, patients receiving granulocyte transfusion therapy usually have been compared to historical controls, despite rapidly changing therapeutic approaches to the malignant diseases in question. Nevertheless, examination of these reports taken all together is quite convincing.

Freireich's group reported the efficacy of CGL cells in neutropenic children. In 12 of 26 instances, defervescence lasting for a median of two days was seen. Five children with *Pseudomonas* septicemia remained afebrile until the development of bone marrow remission one to three weeks later (13). A later report (46) described clinical responses in 54% of 40 patients, again using CGL cells. Interestingly, a temporary bone marrow homograft was detected in one recipient manifested by a change in red cell type. Schwartzenberg reported favorable clinical responses in 50% of 52 leukopenic infected patients receiving CGL cells (47). In addition, he reported graft-versus-host reaction in patients receiving transfusions and noted an association between this syndrome and the ultimate development of remission without chemotherapy in 6 of 24 patients with acute leukemia.

Most recently, Vallejos reported the results of CGL granulocyte transfusions in 128 patients, observing that 49% of serious and life—threatening infections responded to a combination of granulocyte transfusions and antibiotic therapy (48).

Yankee (14) first reported that PMN obtained by multiple phlebotomies from normal donors appeared to function as well as CGL cells clinically and to result in slightly better posttransfusion increments per given dose. However, the difficulties of obtaining and pooling multiple donor buffy coats, and the real and potential risks of alloimmunization and infection made this approach unsatisfactory.

McCredie reported a series of 285 granulocyte transfusions from normal donors given to 73 febrile leukopenic patients. Of 22 recipients receiving between 4 and 8 transfusions, 78% responded with clinical and laboratory evidence of clearing of infection (49). Other investigators (50, 51) have also reported series of cases in which the impression of clinical improvement following transfusions of granulocytes obtained by CFC was noted. Djerassi (52) reported clinical recovery in 20 of 24 consecutively treated neutropenic patients with granulocytes obtained by FL. Herzig, however, in preliminary experiments was unable to show consistent clinical benefit using FL cells, although one patient with *E. coli* septicemia did recover from his infection (53). Evidence of a different sort that granulocyte transfusions are effective was reported by Shohet (54). CGL cells from a female were infused into a male patient with staphylococcal septicemia and neutropenia. Examination of the peripheral blood after transfusion revealed PMN containing Barr bodies which had ingested large numbers of staphylococcal organisms. Concurrently, cytogenetic studies of the peripheral blood revealed the Ph 1 chromosome in all metaphases.

Graw, in an attempt to remedy the lack of adequate controls in such studies, gave 86 transfusions (most from CFC) to 39 neutropenic septicemic patients (55). All transfusions were normal PMN. Thirty seven patients who were being treated concurrently with antibiotics alone were selected from the National Cancer Institute records. Although no significant difference in survival was noted in these groups, all 12 patients who had received 4 or more transfusions survived, while only 5 of 19 control patients who lived long enough to have received 4 or more transfusions recovered from the septicemia (Figure 1).

In the last year, Higby (33) reported the first truly randomized study in which patients who were leukopenic and infected were assigned either to a group receiving antibiotics and four daily granulocyte transfusions (transfused) or to a control group receiving antibiotics alone (not transfused). Of 17 patients in the transfused group, 15 survived to day 20; 13 of the 15 subsequently developed complete remission of disease and one developed partial remission. Of the 19 in the control group, only

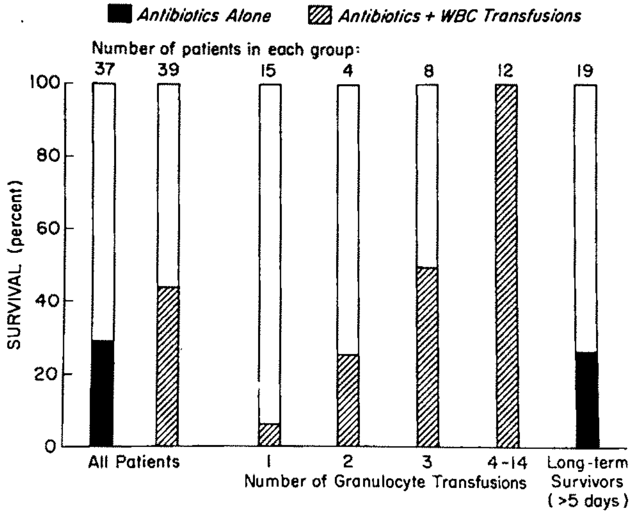


Figure 1 Comparison of survival vs number of transfusions received (reprinted with permission from Ref. 55, p. 369).

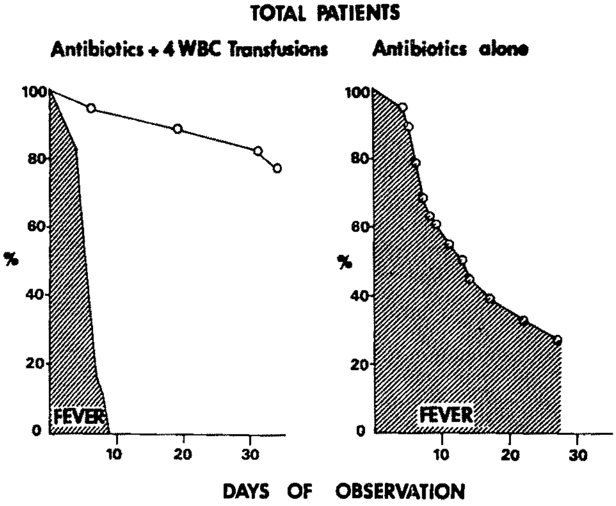


Figure 2 Comparison of patients randomized to receive antibiotics and PMN transfusions vs antibiotics alone.

5 survived to day 20, and 3 of the 5 entered remission (Figure 2). Subgroups of the two groups, based on stage of disease, type and severity of infection, age, and time since chemotherapy had been given, also showed statistically significant improvement in the survival of those receiving granulocyte transfusions.

Normal granulocyte transfusions have also been used in the treatment of other bone marrow disorders. Aplastic anemic patients undergoing marrow grafts sustain prolonged leukopenia, and FL cells have been used for support during this period (56). Raubitschek (57) reported the successful treatment of a child with chronic granulomatous disease complicated by aspergillosis using granulocyte transfusions adjunctive to amphotericin B therapy. Granulocyte transfusions probably will find increasing use in supporting patients with cyclic neutropenia, Chediak-Higashi syndrome, and other disorders related to neutrophil function or number.

In conclusion, granulocyte transfusions have been demonstrated to be therapeutically effective by several investigators in short term treatment of infection in neutropenic patients.

PROCUREMENT OF GRANULOCYTES

Plasmapheresis

Simple plasmapheresis techniques allow the procurement of about 5 X 10^8 granulocytes per unit of blood processed, and the collection of enough normal granulocytes to give a transfusion capable of therapeutic effect requires an inordinate amount of time. The resultant cells, of necessity pooled from several donors, are more likely to result in sensitization and transmission of blood-borne diseases. However, plasmapheresis techniques are still useful in those circumstances in which an appropriate donor with CGL presents at a time when a leukopenic individual requires support. Because of the possibility of engraftment and potentially fatal graft-versus-host disease in these immunosuppressed patients (47), PMN from CGL patients should be irradiated prior to being given in transfusion.

Continuous Flow Centrifugation Leukapheresis

The design of the NCI-IBM cell separator made possible the collection of large numbers of PMN from a single, normal donor (15). In operating the cell separator, antecubital venipunctures are performed on the donor in both arms. Blood is withdrawn from one arm by a roller pump and mixed with an anticoagulant solution automatically. The anticoagulated blood enters the spinning centrifuge bowl. Most of the volume of the centrifuge bowl is displaced by an inner core. Blood separates into red cells lying along the outer wall, leukocytes lying next to the red cell layer, and plasma lying near the inner wall of the cylindrical space. Red cells and plasma are drawn off through their respective ports by separately controlled roller pumps. The operator can allow the buffy coat to accumulate to the point where it is visible and then, by regulating the separate controls, can position the buffy coat beneath the leukocyte port so that a third roller pump can remove the granulocyte-rich fraction. Red cells and plasma are returned to the donor and the leukocytes are collected separately for transfusion.

The technique of CFC leukapheresis can produce approximately 1×10^{10} leukocytes in a single sitting (4–5 hr) from the donor. Limitations in the procedure are expense, donor tolerance, operator fatigue, and the gradual decrease in red cell mass in donors undergoing multiple leukapheresis. Of the white blood cells collected, the fraction of granulocytes varies between 30 and 70%, thus allowing for a granulocyte transfusion of about 5×10^9 cells. Red cell contamination is significant. To increase the total number of granulocytes collected, many modifications of the procedure have been attempted. Various steroids have been used to augment the circulating PMN pool. Etiocholanolone injections result in lasting elevation of granulocyte counts and the use of this agent has been explored extensively by McCredie (17). Prednisolone (29) and dexamethasone (30) have been used to increase the circulating granulocyte count.

The use of endotoxin in this context has largely been abandoned because of the difficulties in establishing a dose-response relationship and because of actual and theoretical dangers associated with the administration of a material capable of activating complement.

The yield of PMN may also be increased by the addition to the inflowing blood of an agent capable of increasing sedimentation of red cells. Dextran, methylcellulose, and fibrinogen have been used in vitro for many years. Dextran has been used by Goldman to augment granulocyte separation in humans (58). Benbunan has used modified fluid gelatin (51). A new agent, hydroxyethyl starch, does not appear to be allergenic or otherwise hazardous and is readily metabolized in vivo (59). Various investigators (17, 29, 30) have been able to increase the yield of PMN tenfold by using hydroxyethyl starch alone or with steroids.

Complications of donation of granulocytes by means of CFC leukapheresis are rare. Chills and fever have been reported on occasion (16). As with any procedure involving venipuncture and withdrawal of blood, occasional fainting spells are seen. Phlebitis and venous thrombosis can occur. Bleeding has not been a problem in our experience, even with individuals with severe thrombocytopenia (as when obtaining circulating leukemic cells for experimental purposes). Graft-versus-host disease has occurred following transfusion of normal leukocytes from CFC (80) but is very rare. As with CGL cells, CFC cells should be irradiated prior to transfusion to prevent this possibility.

CFC methods can be used to accumulate extremely large numbers of cells from CGL patients. Treatment of CGL by multiple CFC leukapheresis has met with varying degrees of success (60, 61).

Filtration Leukapheresis

The phenomenon of leukoadhesion (the adherence of granulocytes to "active" foreign substances) was first applied practically for laboratory purposes to obtain blood lymphocytes relatively uncontaminated by PMN. Glass wool columns were used (62). Djerassi (18) described a system using a commercially available nylon wool pack (Fenwal's Leukopak ®). When whole blood is introduced into these "filters," PMN and monocytes selectively adhere to the nylon fibers while other blood components pass through and can be returned to the donor. Djerassi refined this technique in 1972 (63) and since that time he and his associates have designed an automated pressure pump which permits rapid processing of 16 to 32 units of donor blood in 2–4 hr. Following the procedure, the cells are eluted from the nylon wool using a solution containing a calcium chelating agent (usually ACD solution). The efficiency of elution can be increased if plasma or a solution containing plasma is used as a carrier (53, 64). Varying reports on the in vitro functions of the cells obtained by FL exist. Herzig et al (53) showed a decrease (compared to cells obtained directly from peripheral blood) of 6% in viability as determined by phase microscopy, a decrease of 9% in terms of radiolabeled phagocytosis and a decrease of 11.7% in terms of bactericidal capacity. Similar in vitro results were noted by Higby et al (33). More recent work by Djerassi suggests that cells can be obtained by FL which appear to be identical to normal cells in all measured parameters (52). The yield of cells obtained by FL can be as high as 1×10^{11} granulocytes from a normal donor (52). However, most investigators report yields of approximately $2–5 \times 10^{10}$ granulocytes in 2–3 hr (33, 64, 41). (Thus, FL techniques are 5–20 times as effective as CFC in procuring PMN per given period of time.) Augmentation of peripheral granulocyte counts with dexamethasone prior to the procedure results in improvement in yield by 20% (33). However, the filtration process itself causes leukocytosis (53) which is preceded by profound but very brief neutropenia (41).

At Roswell Park, in over 500 donations, no morbidity of consequence was seen. Seventy consecutive donors were studied in detail. All donors received 6 mg of dexamethasone i.v. and 17,500 units of heparin during the procedure. Donors showed a mean drop of 0.8 g% of hemoglobin and of $30{,}000/\text{mm}^3$ of platelets immediately following the procedure, but both returned to normal by the following day in every instance. The mean increase in granulocytes per mm^3 was 3900. Selected donors were studied for changes in serum hemoglobin, serum haptoglobin, fibrinogen, fibrin split products, and changes in morphology of red cells, white cells, and platelets. No abnormalities were noted. Clotting times in the donors invariably returned to normal within 2 or 3 hr following termination of the procedure.

One donor underwent leukapheresis three times per week for four weeks. Initial values were: hemoglobin 15.4 g%, granulocyte count 4,300/mm^3, and platelet count 450,000/mm^3. The values on the day following the last procedure were: hemoglobin 13.0 g%, total granulocyte count 5600/mm^3, and platelet count 90,000/mm^3. Serum protein electrophoresis, serum haptoglobin, and serum hemoglobin were not changed significantly from initial values.

Faintness occurred in 21% of donors and two actually fainted; chills occurred once. Two of more than 500 donors developed severe abdominal cramps which rapidly resolved (33).

Recipients of FL-produced cells frequently develop chills and fever during and following the transfusion (33, 41, 55). Schiffer et al (41) observed severe toxic reactions in nine patients on 18 occasions, characterized by hyperpyrexia, occasional hypotension, dyspnea, and wheezing. In all, the overall incidence of significant reactions was 27% as compared to an incidence of about 12% with CFC cells. Djerassi (52) and Higby (33), however, have seen no serious reactions to granulocyte transfusions obtained by FL. Most likely, methods of procuring granulocytes for transfusion and rates of administration bear on the response seen in the recipient.

In early studies at Roswell Park, rapid infusion of cells resulted in dyspnea, wheezing, cyanosis, and pulmonary infiltrates in two patients. Because of this, rate of administration was modified to resemble the method used by Djerassi. Cells were administered at the rate of 10^{10} per 20 to 30 min, with a 30-min rest following each aliquot. Thus, a typical transfusion of granulocytes may take up to 4–6 hr. Since instituting this approach, no serious reactions have been seen in over 500 transfusions. (It should be noted that this does not appear to be a problem with CFC leukapheresis. This may be because in CFC, unlike FL, the cells are not subjected to a membrane-altering stress prior to infusion.) Although 1-hr posttransfusion counts of PMN obtained by FL are probably not comparable to those obtained in CFC leukapheresis due to differing methods of administration and differing kinetics of in vivo survival (53), Graw reported a corrected PMN increment of 850 (range 0–3900) in 68 transfusions using the Blood Cell Separator; whereas, the increment seen with the filtration leukapheresis system was 233, despite the administration of cell doses four times larger on the average in these 28 transfusions (55). Other investigators (33, 52) described posttransfusion increments ranging from 50 to 3000 granulocytes/mm^3 10^{10} cells given per M^2 surface area. Since 95% of white cells obtained by FL are PMN, it is unnecessary to irradiate them prior to infusion.

DONOR-RECIPIENT COMPATIBILITY

The importance of matching donor and recipient according to blood group type, HL-A type, or granulocyte type is not fully known. Posttransfusion increments may not be directly related to clinical outcome. The severity of infection in the recipient and the granulocyte reserve present (which may not be reflected in peripheral white blood cell counts) almost certainly are important factors in determining posttransfusion counts.

In Graw's study (55) 1-hr posttransfusion counts (using granulocytes obtained by CFC) were greatly reduced with ABO group discordance. On the other hand, Vallejos stated that the 1-hr posttransfusion increment was similar in all patients regardless of ABO group (48). Probably, ABO group compatibility testing is important in direct proportion to the amount of red blood cells given in a white cell transfusion, and perhaps in the degree of pre-existing alloimmunization of the recipient to red cell antigens.

Posttransfusion increments are much more closely related to HL-A compatibility. The relationship between posttransfusion count and degree of relatedness between donor and recipient is directly proportional (50, 65). With CFC cells, the posttransfusion increment is inversely proportional to the number of HL-A antigens not held in common by donor and recipient. Thus, if there are no mismatches, the increment is higher than if there is one, etc (55, 66) (Figure 3). At this writing, the relationship between posttransfusion increments and HL-A compatibility is not clear with respect to FL-obtained cells.

Despite this relationship, it appears that, at least over the short term, HL-A compatibility and posttransfusion increments are not serious determining factors in clinical outcome, either for CFC cells (67) or FL cells (33). No clinical studies have been done comparing the effects of HL-A matched versus unmatched cells, and it may be that it would require fewer matched cells than unmatched cells to resolve an infection. It is expected that HL-A matched cells would be most important if long term support (as in aplastic anemia) were necessary.

Antileukocyte antibodies in the recipient are probably of more importance in relation to the clinical outcome of granulocyte transfusions. The presence of leukocyte antibody impaired the ability of granulocytes to kill *Staphylococcus* in vitro. One-hour posttransfusion increments following granulocyte transfusion are markedly reduced in the presence of leukocyte antibody (68). Labeled granulocytes transfused into recipients with leukocyte antibodies are sequestered in the spleen, while in recipients without antibodies, localization in areas of infection can be demonstrated (69). Despite this, Hester was unable to show a relationship between leukoagglutinins in the recipients and clinical response (67).

Leukoagglutinating antibodies are associated with hazards to the recipient also. Ward (70) described, in patients with leukoagglutinins receiving transfusions, a syndrome characterized by chills, fever, tachycardia, dyspnea, and coughing, with roentgenologically demonstrable pulmonary infiltrates. Eosinophilia is commonly present. The syndrome is usually self-limited and is not predictable on the basis of a cross-match. Similar situations have been noted by Schiffer (41) and by Higby (33) with granulocyte transfusions. In view of the rapidity of onset of this syndrome, its cause is probably a hypersensitivity reaction. Prediction of this complication in an individual donor-recipient setting is not yet possible.

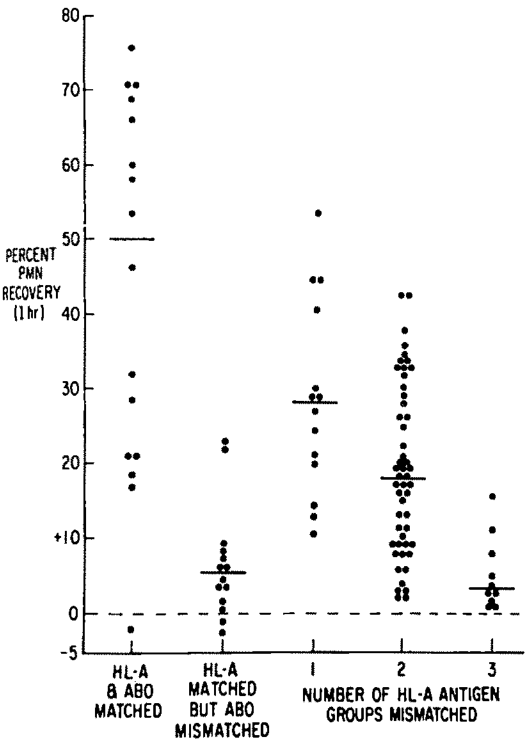


Figure 3 The relationship between transfused PMN's at 1 hr and the number of HL-A mismatches (reprinted with permission from Ref. 55, p. 368).

Separate and distinct antigens specific to granulocytes have been described by Lalezari (71). Their usefulness in granulocyte transfusions has not yet been determined. At the present time, minimal requirements for granulocyte transfusion should be the absence of reaction in a donor—recipient cross-match. Great caution should be exercised if the recipient has preformed leukocyte antibodies to donor leukocytes. ABO and HL-A matching should be as close as possible under the circumstances available, but absence of matched donors should not inhibit the therapist. Exploration of the importance of HL-A typing, granulocyte typing, and leukoagglutinin reactions may require multi-institutional cooperative studies.

FUTURE WORK

Mode of Administration

Currently granulocyte transfusions are given as often as possible to recipients having leukopenia and fever. Daily administration may not be optimal, however, and clinical investigation based on known kinetics of transfused granulocytes must be performed. Freireich et al showed that the number of granulocytes given were proportional to the percentage of patients in each group becoming afebrile (72). Thus, it may be more advantageous to administer extremely large doses of granulocytes in one dose. On the other hand, frequent infusions of lower doses of granulocytes may be advantageous if the in vivo survival and/or duration of action of the cells is short, regardless of the number of granulocytes administered.

Prophylactic Use

Whether or not granulocytes should be administered prophylactically remains to be shown. Morbidity in leukopenic patients could be decreased and the investment in granulocyte procurement reduced by giving smaller doses of granulocytes at infrequent intervals prior to the development of fever. Cooper et al (73) are conducting a study in which neutropenic and thrombocytopenic patients are randomized to receive, prophylactically, either transfusions of granulocytes and platelets or platelets alone. The study so far suggests a decreased incidence of infection in the group receiving granulocytes and platelets.

Preservation of Granulocytes

Granulocytes represent the end stage of maturation and are highly differentiated to undergo cytolysis in response to various stimuli. Macroagglutination (74) and lysozymal injury (75) interfere with the recovery of usable granulocytes following cryopreservation. The development of granulocyte banks, a necessary adjunct to widespread use of megadoses of chemotherapy, awaits progress in this field. Crawley & Valery demonstrated the preservation of some function in granulocytes subjected to cryopreservation (76) and Cavins et al were able to preserve granulocyte function in 90% of cells cryopreserved with DMSO, but the lots of cells used were small (77). Although granulocyte cryopreservation has not yet become a practical reality, it would allow for the preservation of autologous cells from patients in remission, thereby allowing intensive consolidation therapy with marrow ablative doses of chemotherapy. Stockpiling of autologous cells from patients in remission, especially those with acute leukemia, would allow for safer remission induction following relapse.

In Vitro Production of Granulocytes

In the future granulocytes may be produced in culture systems from stem cells obtained from normal donors. Maturation of normal myeloblasts in vitro has been demonstrated (78). Gailani, at Roswell Park, has recently demonstrated that marrow tissue can be maintained in organ culture and continue to produce PMN and other elements for long periods of time (79).

CONCLUSIONS

Interest in granulocyte transfusion therapy has increased explosively concurrent with the development of more aggressive chemotherapeutic and radiotherapeutic treatment of cancer and related disorders. Demonstration that this therapy can reverse the consequences of infection in neutropenic individuals has come from several sources. Normal granulocytes can be procured relatively easily and the technical problems associated with transfusion of granulocytes have been, for the most part, overcome. Investigation needs to be done in several areas, including the role of tissue typing, the frequency and mode of administration of granulocytes, the prophylactic and therapeutic use of such transfusions, and the preservation of granulocytes. The use of granulocyte transfusions opens up new fields for the chemotherapist. The possibility of sustaining patients through bone marrow transplantation procedures is greatly increased. Clinical investigation of total body irradiation can now be carried out more safely. The limitations to chemotherapy imposed by anemia, granulocytopenia, and thrombocytopenia are no longer absolute, thus allowing for development of more intensive regimens. At this time granulocyte transfusions should be available to every patient undergoing marrow ablative therapy, especially those with acute leukemia.

ACKNOWLEDGMENT

The authors wish to thank Mrs. Wende Kurk for her patience and skill in preparing this manuscript.

Literature Cited

1. Levine, A. S., Schimpff, S. C., Graw, R. G. Jr., Young, R. C. 1974. Hematologic malignancies and other marrow failure states: Progress in the management of complicating infections. *Semin. Hematol.* 11:141–202
2. Feld, R., Bodey, G. P., Rodriguez, V. 1974. Causes of death in patients with malignant lymphoma. *Proc. AACR and ASCO, March 1974, Houston, Texas,* Abstr. 807, p. 185
3. Bodey, G. P. 1973. Infections in patients with cancer. In *Cancer Medicine,* ed. J. F. Holland, E. Frei, 1135–65. Philadelphia: Lea and Febiger
4. Bloomfield, C. D., Kennedy, B. J. 1974. Cepalothin, carbenicillin, and gentamicin combination therapy for febrile patients with acute non-lymphocytic leukemia. *Cancer.* 34:431–37
5. Schimpff, S., Satterlee, W., Young, V. M., Serpick, A. 1971. Empiric therapy with carbenicillin-gentamicin. *N. Engl. J. Med.* 284:1061–65
6. Tattersall, M. H. N., Spiers, A. S. D., Darrell, J. H. 1972. Initial therapy with combination of five antibiotics in febrile patients with leukemia and neutropenia. *Lancet* 1:162–65
7. Bodey, G. P., Gehan, F. A., Freireich, E. J., Frei, E. 1971. Protected environment prophylactic antibiotic program in the chemotherapy of acute leukemia. *Am. J. Med. Sci.* 262:138–51
8. Levine, A. S. et al 1973. Protected environments and prophylactic antibiotics: A prospective controlled study of their utility in the therapy of acute leukemia. *N. Engl. J. Med.* 288:477–83
9. Yates, J. W., Holland, J. F. 1973. A controlled study of isolation and endogenous microbial suppression in acute myelocytic leukemia patients. *Cancer* 32:1490–98
10. Bodey, G. P., Buckley, M., Sathe, Y. S., Freireich, E. J. 1966. Quantitative relationships between circulating leukocytes and infection in patients with acute leukemia. *Ann. Intern. Med.* 64:328–36
11. Strumia, M. M. 1934. The effect of leukocyte cream injections in the treatment of the neutropenias. *Am. J. Med. Sci.* 187:527–44
12. Brecher, G., Wilbur, K. M., Cronkite, E. P. 1953. Transfusion of separated leukocytes into irradiated dogs with aplastic marrows. *Proc. Soc. Exp. Biol. Med.* 84:54–56
13. Morse, E. E., Bronson, W., Carbone, P. P., Freireich, E. J. 1961. Effectiveness of granulocyte transfusion from donors with chronic myelocytic leukemia to patients with leukopenia. *Clin. Res.* 9:332
14. Yankee, R. A., Freireich, E. J., Carbone, P. P., Frei, E. 1964. Replacement therapy using normal and chronic mye-

logenous leukemia leukocytes. *Blood* 24:844–45
15. Freireich, E. J., Judson, G., Levin, R. H. 1965. Separation and collection of leukocytes. *Cancer Res.* 25:1516–20
16. Buckner, D., Graw, R. G., Eisel, R. J., Henderson, E. S., Perry, S. 1969. Leukapheresis by continuous flow centrifugation in patients with chronic myelocytic leukemia. *Blood* 33:353–69
17. McCredie, K. B., Freireich, E. J. 1971. Increased granulocyte collection from normal donors with increased granulocyte recovery following transfusion. *Proc. Am. Assoc. Cancer Res.* Abstr. 232, p. 58
18. Djerassi, I., Kim, J. S., Mitrakul, C., Suvansri, U., Ciesileka, W. 1970. Filtration leukapheresis for separation and concentration of transfusable amounts of normal human granulocytes. *J. Med. (Basel)* 1:358–64
19. Henderson, E. S. 1974. The granulocytopenic effects of cancer chemotherapeutic agents. In *Drugs and Hematologic Reactions, The Twenty-Nineth Hahnemann Symposium,* ed. N. V. Dimitrov, J. H. Nodine, 207–22. New York: Grune & Stratton
20. Cartwright, G. E., Athens, J. W., Wintrobe, M. M. 1964. The kinetics of granulopoiesis in normal man. *Blood* 24:780–803
21. Craddock, C. G., Perry, S., Lawrence, J. S. 1960. The dynamics of leukopenia and leukocytosis. *Ann. Intern. Med.* 52:281–93
22. Marsh, J. C., Boggs, D. R., Cartwright, C. E., Wintrobe, M. M. 1967. Neutrophil kinetics in acute infection. *J. Clin. Invest.* 46:1943–52
23. Boggs, D. R., Chervanik, P. A. 1972. Chemotherapy and leukokinetics. In *Cancer Chemotherapy II, Twenty-second Hahnemann Symposium,* ed I. Brodsky B. Kahn. New York: Grune & Stratton
24. Boggs, D. R. 1967. The kinetics of neutrophilic leukocytes in health and disease. *Semin. Hematol.* 4:359–86
25. Kauder, E. et al 1965. Leukokinetic studies, XII. Kinetic studies of normal isologous neutrophilic granulocytes transfused into normal subjects. *Proc. Soc. Exp. Biol. Med.* 120:595–97
26. Brittingham, T. E., Chaplin, H. 1957. Febrile transfusion reactions caused by sensitivity to donor leukocytes and platelets. *J. Am. Med. Assoc.* 165:819–25
27. Herzig, R. H., Poplack, D. G., Yankee, R. A. 1974. Prolonged granulocytopenia from incompatible platelet transfusions. *N. Engl. J. Med.* 290:1220–23
28. Samuels, A. J. 1951. Primary and secondary leukocyte changes following the intramuscular injection of epinephrine hydrochloride. *J. Clin. Invest.* 30:941–47
29. Pfister, H., Ruppert, W. 1972. High effective technique for procurement of viable granulocytes with the IBM blood cell separator. *Am. Assoc. Blood Banks Ann. Meet., Washington, August, 1972*
30. Mishler, J. M., Higby, D. J., Rhomberg, W., Holland, J. F., Nicora, R. W. 1974. Hydroxyethyl starch and dexamethasone as an adjunct to leukocyte separation with the IBM blood cell separator. *Transfusion* 14:352–56
31. Bishop, C. R. et al 1968. Leukokinetic studies XIII. A non-steady state kinetic evaluation of the mechanism of cortisone-induced granulocytosis. *J. Clin. Invest.* 47:249
32. Fruhman, G. J. 1962. Adrenal steroids and neutrophil mobilization. *Blood* 20:355–62
33. Higby, D. J., Henderson, E. S., Holland, J. F. 1974. Granulocyte transfusion therapy: Clinical and laboratory studies. *Proc. AACR and ASCO, March 1974, Houston, Texas,* Abstr. 771
34. Vogel, J. M., Yankee, R. A., Kimball, H. R., Wolff, S. M., Perry, S. 1967. The effect of etiocholanalone on granulocyte kinetics. *Blood* 30:474–84
35. Vogel, J. M., Kimball, H. R., Wolff, S. M., Perry, S. 1967. Etiocholanalone in the evaluation of marrow reserves in patients receiving cytotoxic agents. *Ann. Intern. Med.* 67:1226–38
36. Cline, M. J., Melmon, K. L., Davis, W. C., Williams, H. E. 1968. Mechanism of endotoxin interaction with human leucocytes. *Brit. J. Hematol.* 15:539–47
37. Boggs, D. R., March, J. C., Chervenick, P. A., Cartwright, G. E., Wintrobe, M. M. 1968. Neutrophil releasing activity in plasma of normal human subjects injected with endotoxin. *Proc. Soc. Exp. Biol. Med.* 127:689–93
38. Jensen, D. P., Brubaker, L. H., Nolph, K. D., Johnson, C. A., Nothum, R. J. 1973. Hemodialysis coil-induced transient neutropenia and overshoot neutrophilia in normal man. *Blood* 41:399–408
39. Brubaker, L. H., Nolph, K. D. 1971. Mechanisms of recovery from neutro-

penia induced by hemodialysis. *Blood* 38:623–31

40. Toren, M., Goffinet, A., Kaplow, L. S. 1970. Pulmonary bed sequestration of neutrophils during hemodialysis. *Blood* 36:337–40
41. Schiffer, C. A., Bucholz, D. H., Alsner, J., Betts, S. W., Wiernik, P. 1974. Clinical experience with transfusion of granulocytes obtained by continuous-flow filtration leukapheresis. *Proc. AACR and ASCO, March 1974, Houston, Texas,* Abstr. 103
42. Weisberger, A. S., Heinle, R. W., Hannah, R. 1949. Transfusion of leukocytes and products of disintegrated leukocytes. *Proc. Soc. Exp. Biol. Med.* 70:749
43. Epstein, R. B., Clift, R. A., Thomas, E. D. 1969. The effect of leukocyte transfusions on experimental bacteremia in the dog. *Blood* 34:782–80
44. Debelak, K. M., Epstein, R. B., Anderson, B. R. 1974. Granulocyte transfusion in leukopenic dogs: In vivo and in vitro function of granulocytes obtained by continuous flow filtration leukopheresis. *Blood* 43:757–66
45. Dale, D. C. et al 1974. Granulocyte transfusion therapy of experimental pseudomonas pneumonia. *J. Clin. Invest.* 54:664–71
46. Morse, E. E., Freireich, E. J., Carbone, P. P., Bronson, W., Frei, E. 1966. The transfusion of leukocytes from donors with chronic myelocytic leukemia to patients with leukopenia. *Transfusion* 6:183–92
47. Schwartzenberg, L. et al 1967. Study of factors determining the usefulness and complications of leukocyte transfusion. *Am. J. Med.* 43:206–13
48. Vallejos, C., McCredie, K. B., Bodey, G. P., Hester, J. P., Freireich, E. J. 1975. White blood cell transfusions for control of infections in neutropenic patients. *Transfusion* 15:28–33
49. McCredie, K. B., Vallejos, C. S., Hester, J. P., Freireich, E. J. 1974. The results of granulocyte transfusions collected from normal donors. *Proc. AACR, March 1974, Houston, Texas,* Abstr. 480, p. 120
50. Koza, I., Holland, J. F., Cohen, E. 1971. Histocompatible leukocyte transfusion during granulocytopenia. *Neoplasma* 18:185–91
51. Benbunan, M., Bussel, A., Reviron, J., Boiron, M., Bernard, J. 1973. Les transfusions de granulocytes normaux preleves sur separateur de cellules. Resultats techniques et cliniques. *Nouv. Rev. Fr. Hematol.* 13:469–86
52. Djerassi, I., Kim, J. S., Survansri, U. 1973. *Filtration Leukopheresis: Technique of Filtration, Properties of Filtered Granulocytes and Preliminary Clinical Studies on Transfusions of Filter-Collected Leukocytes.* Presented at the Leukemia Center (Platelet) Program Conf., Dec. 10, 1973. NIH, Bethesda, Maryland
53. Herzig, G. P., Root, R. K., Graw, R. G. 1972. Granulocyte collection by continuous flow filtration leukopheresis. *Blood* 39:554–67
54. Shohet, S. B. 1968. Morphologic evidence for the in Vivo activity of transfused chronic myelogenous leukemia cells in a case of massive staphylococcal septicemia *Blood* 32:111–18
55. Graw, R. G., Herzig, G., Perry, S., Henderson, E. S. 1972. Normal granulocyte transfusion therapy: Treatment of septicemia due to gram negative bacteria. *N. Engl. J. Med.* 287:367–71
56. Storb, R. et al 1971. Allogenic marrow grafting for treatment of aplastic anemia. *Blood* 43:157–80
57. Raubitschek, A. A., Levin, A. S., Stites, D. P., Shaw, E. B., Fudenberg, H. H. 1973. Normal granulocyte infusion therapy for aspergillosis in chronic granulomatous disease. *Pediatrics* 51:230–33
58. Goldman, J. M. Personal communication
59. Thompson, W. L., Walton, R. F. 1962. Parenteral administration of hydroxyethyl starches. *Nat. Acad. Sci. Conf. Artificial Colloids for Intravenous Use,* pp. 168–206
60. Vallejos, C. S., McCredie, K. B., Brittin, G. M., Freireich, E. J. 1973. Biological effects of repeated leukapheresis of patients with chronic myelogenous leukemia. *Blood* 42:925–33
61. Hadlock, D. C., McCullough, J. J., Dienard, N., Kennedy, B. J., Fortuny, I. E. 1974. Role of continuous-flow centrifuge leukapheresis in the management of chronic myelogenous leukemia. *Proc. AACR and ASCO, March 1974, Houston, Texas*
62. Johnson, T. M., Garvin, J. E. 1959. Separation of lymphocytes in human blood by means of glass wool columns. *Proc. Soc. Exp. Biol. Med.* 102:333–35
63. Djerassi, I., Kim, J. S., Suvansri, U., Mitrakul, C., Ciesielka, W. 1972. Continuous flow filtration-leukapheresis. *Transfusion,* 12:75–83

64. Hill, Norwood, O. Personal communication
65. McCredie, K. B., Freireich, E. J., Hester, J. P., Vallejos, C. 1973. Leukocyte transfusion therapy for patients with host-defense failure. *Transplantation Proc.* 3:1285–89
66. Higby, D. J. et al 1974. Increased elevation of peripheral leukocyte counts by infusion of histocompatible granulocytes. *Vox Sang.* 27:186–90
67. Hester, J. P., Rossen, R. O. 1974. Multiple granulocyte transfusions: Role of HL-A compatibility and leukoagglutinins. *Proc. AACR and ASCO March 1974, Houston, Texas*
68. Goldstein, I. M., Eyre, H. J., Terasaki, P. I., Henderson, E. S., Graw, R. G. 1971. Leukocyte transfusions: Role of leukocyte alloantibodies in determining transfusion response. *Transfusion* 11:19–24
69. Eyre, H. J., Goldstein, I. M., Perry, S., Graw, R. G. 1970. Leukocyte transfusions: Function of transfused granulocytes from donors with chronic myelocytic leukemia. *Blood* 36:432–41
70. Ward, H. N. 1970. Pulmonary infiltrates associated with leukoagglutinin transfusion reactions. *Ann. Intern. Med.* 73:689–94
71. Lalezari, P., Bernard, G. E. 1966. An isologous antigen-antibody reaction with human neutrophiles, related to neonatal neutropenia. *J. Clin. Invest.* 45:1741–50
72. Freireich, E. J. et al 1964. The function and fate of transfused leukocytes from donors with chronic myelocytic leukemia in leukopenic recipients. *Ann. N.Y. Acad. Sci.* 113:1081–89
73. Cooper, M. R., Pajak, T., Richards, F., Burroghs, R., Spurr, L. 1973. *Histocompatible Leukocyte and Platelet Transfusions.* Presented at Leukemia Center Program Conf., Dec. 10, 1973, Bethesda, Maryland
74. Bock, G. N., Chess, L., Mardina, M. R. 1972. The prevention of clumping of frozen stored leukocyte populations by EDTA. *Cryobiology* 9:216–18
75. Lotke, P. A. 1966. Lysozyme stabilizing agents for hypothermic kidney preservation. *Nature* 212:512–13
76. Crawley, J. P., Valery, C. R. 1973. The recovery, viability and function of leukocytes freeze-preserved in dimethylsulfoxide. *Soc. Cryobiol., 10th Ann. Meet.* Abstr. 57
77. Cavins, J., Djerassi, I., Aghai, E., Roy, A. J. 1968. Current methods for the cryopreservation of human leukocytes (granulocytes). *Cryobiology* 5:60–69
78. Chervanik, P. A., Boggs, D. R. 1971. In vitro growth of granulocytic and mononuclear cell colonies from blood of normal individuals. *Blood* 37:131–35
79. Gailani, S., Nussbaum, A., Silvernail, P., Roholt, O. 1974. Studies on production of an osteolytic factor by myeloma cell culture. *Proc. AACR and ASCO, March 1974, Houston, Texas,* Abstr. 405

Copyright 1975. All rights reserved

ACUTE RENAL FAILURE (VASOMOTOR NEPHROPATHY): MICROPUNCTURE STUDIES OF THE PATHOGENETIC MECHANISMS

❖7152

Donald E. Oken, M.D.[1]
Department of Medicine, Peter Bent Brigham Hospital and Harvard Medical School,
Boston, Massachusetts 02115

Both the pathoanatomic and clinical consequences of mismatched blood transfusions were well known forty years ago (1), and pathologic alterations in kidneys of World War I battle casualties were carefully documented (2, 3). A report of two patients who developed acute renal failure resulting from trauma was published in 1937 (4), but the relationship between shock, trauma, and acute oliguric renal failure first became widely appreciated in 1940. Then, with large numbers of casualties surviving severe war wounds, the clinicopathologic correlation quickly became evident and the natural history of this form of nephropathy was established. It was soon discovered that an identical syndrome could occur after hemolytic anemia, blackwater fever, heat stroke, sepsis, surgical complications, or poisoning with mushrooms, carbon tetrachloride, and a host of other noxious agents.

As the frequency of this abnormality and its high death toll were recognized, it became a priority research item. Numerous physiologic studies were undertaken (5–7), but once it was recognized that the markers used for clearance determinations might leak across damaged tubular epithelium, the validity of such measurements became questioned. Concepts of pathogenesis, then, stemmed largely by inference from the three major pathologic abnormalities observed: tubular necrosis, intraluminal "casts," and interstitial edema. Tubular necrosis gave rise to the "passive backflow" thesis while intraluminal debris and interstitial edema suggested the obstructive theory. A vascular etiology assumed by Hackradt (8) as early as 1917 was neglected despite clear evidence of renal ischemia (9) obtained by Trueta and

---

[1] Supported by U.S. Public Health Service Grant AM 08695.

his colleagues. The important hemodynamic aberrations in human and experimental acute renal failure have been summarized elsewhere (10–12) and cannot be included in this brief review.

According to the passive backflow theory, glomerular filtrate is formed at a normal, or only slightly reduced, rate and is quantitatively reabsorbed across injured epithelium which can no longer serve as an effective absorption barrier. All components of the filtrate—creatinine, glucose, sodium, water, etc—are postulated to leak out of areas of injured epithelium. Of the classical studies frequently considered as evidence for this theory (13–20), not one, in fact, actually proved the existence of passive backflow. As pointed out in a recent review (21), backflow has been assumed in an effort to explain findings that could equally well have been ascribed to filtration failure and depressed active transport by injured tubule epithelium.

The tubular obstructive theory presupposed that intraluminal debris and proteinaceous material found on histologic sections obstruct urine outflow, raising intratubular pressure and preventing further filtration. Increased intraluminal pressure also would tend to augment any tendency to passive backflow. Filtration is supposed to continue at a high rate according to the backflow theory and to occur only at a rate equivalent to absorption in the obstructive theory. The two phenomena, theoretically, might co-exist, however. If passive absorption were sufficiently rapid proximal to a site of obstruction, any tendency to increased intratubular pressure would be dissipated and filtration could be well maintained.

Until the application of renal micropuncture to this problem some 12 years ago, the exact role played by each possible functional aberration was purely conjectural. This review is written to examine the results of micropuncture studies which have, in this author's opinion, now clearly characterized the mechanisms responsible for many models of experimental acute renal failure. The assumption that any experimental model is a meaningful analogue of human vasomotor nephropathy remains to be established, however. Certain models closely resemble the human disease histologically, in the functional circumstances leading to renal failure, and the reversible course of renal insufficiency—glycerol-induced myohemoglobinuric renal failure, for example. With plasma volume depletion, release of myoglobin from damaged muscle and hemolysis, it has many features of the so-called crush syndrome. Poisoning with nephrotoxins such as $HgCl_2$, usually with suicidal intent, is directly reproducible in the laboratory and, interestingly, shows similar functional aberrations to those of glycerol-induced and methemoglobinuric renal failure. One would not, on the other hand, assume that total ureteral occlusion, experimental glomerulonephritis, and hypersensitivity angitis have the same pathogenesis as vasomotor nephropathy or each other. The administration of doses of folic acid which approach the limit of its aqueous solubility inevitably results in renal failure due to tubular blockade (22). Micropuncture study shows abnormalities after such "megadosage" of folic acid which are very different from those demonstrable in most other models (22). It would be importune, indeed, if one were to generalize about the etiology of acute renal failure on the basis of this model, tempting though it might be. Similarly, total renal arterial occlusion of hours' duration presents an extreme degree of renal ischemia which is unlikely to be matched under most circumstances in man. While renal perfusion is reduced under conditions of shock, hemorrhage, trauma, and sepsis, the kidney is not completely ischemic. Prolonged total ischemia of any tissue results in its death. Extrapolation from studies of a dead or inevitably dying tissue to the circumstances of an injured but potentially recoverable one is, to say the least, questionable. Studies on folic acid nephropathy (22) and total renal ischemia (23), although of interest in their own right, have, then, been omitted from this review.

Renal micropuncture affords the best means presently available for evaluating the various theories of the pathogenesis of acute renal failure, particularly when renal blood flow and its compartmental distribution are studied concomitantly. One can derive the glomerular filtration rate (GFR) and fractional water absorption in single nephrons from the tubule fluid:plasma inulin concentration ratio, measure intratubular pressures by direct manometry, and assess tubule fluid flow rates from the transit time of intravenously injected lissamine green. In using inulin to quantify filtration and fluid absorption, one must be confident that significant amounts of this marker have not been absorbed. Certain observations can be made, however, which make no assumptions as to the tubules' permeability to inulin. These qualitative observations are direct and give irrefutable results. Observing the surface nephrons, for example, one can see whether tubules are full of fluid or are collapsed; injecting a small drop of stained mineral oil, one can clearly discern whether fluid in a given tubule flows briskly, slowly, or not at all; the rate of absorption of a segment of isotonic saline injected between two oil droplets is readily appreciable to the eye if it is grossly subnormal. Such simple observations can afford very valuable information. Under steady state conditions, any fluid which has not been absorbed up to a given site in a tubule will be propelled onward by new filtrate reaching that site. If flow in a fluid-filled tubule is very sluggish or totally absent, filtration in that tubule must be scant indeed. In addition, it is hard to see how fluid absorption can be anything but slight if nephrons remain fluid-filled and, in the absence of flow, little fresh filtrate is available to replace that which has been absorbed.

The first micropuncture study of the pathophysiology of acute renal failure was performed on rats poisoned with 12 mg/kg body weight (BW) $HgCl_2$ (24). Single nephron GFR (SNGFR), tubule fluid flow rate, proximal tubule sodium and water absorption, and intratubular pressures were measured in individual proximal tubules during both the developmental and established phases of renal insufficiency. Within 6 hr of injection, SNGFR had fallen to one half of control but intratubular pressure had remained normal. So long as flow was brisk, debris flowed down the tubule with no apparent tendency to aggregate. As flow progressively decreased, however, the debris sometimes coalesced into potentially occlusive masses within the tubule lumen. By this time, intratubular pressure was decidedly subnormal so that the fall in GFR could not be attributed to the presence of this accreted material and increased pressure. Instead, it appeared that any element of obstruction which developed was actually the result of the decrease in GFR, greatly decreased flow permitting accumulation of cell debris which might have been propelled out of the tubule otherwise. While filtration was sufficiently brisk to permit accurate volume measurement, proximal fractional water absorption was essentially normal. When filtration had essentially stopped, fractional water absorption, estimated from the rate of absorption of a segment of 140-mM saline injected between oil droplets, was grossly reduced. All indications pointed to decreased, rather than increased, fluid absorption from the proximal tubule. It was suggested, therefore, that impaired glomerular filtration is the prime abnormality leading to the development of this form of acute renal failure in the rat, that obstruction was a secondary phenomenon, and that evidence of passive backflow could not be elicited.

Bank et al (25) performed renal micropuncture studies which strongly supported the passive backflow concept. Using animals injected 24 hr earlier with 4 to 4.7 mg/kg BW $HgCl_2$, they found that intravenously injected lissamine green was filtered in many nephrons but became progressively more pale and failed to appear at all in distal tubules. Since lissamine green normally becomes more concentrated in color as it traverses the proximal tubule and densely stains the fluid in distal tubule segments, it appeared as if the dye had leaked away across damaged proximal tubule epithelium. The inulin clearances measured in the early proximal tubule were not different from control; clearances measured in the midportion of the proximal tubule were only 37% of control. This difference was interpreted, quite logically, as an indication that inulin had been lost from the tubule in much the same way as lissamine green. Such a finding was quite remarkable, however, since the proximal two thirds of the proximal tubule does not show necrosis after the dose of mercuric chloride used (26). It is particularly noteworthy since only 29% of the remaining inulin was said to have been lost from the distal one third of the proximal tubule, the only segment of the nephron displaying frank tubular necrosis.

Steinhausen and co-workers (27) repeated the study of Bank et al with a comparable dose of mercuric chloride. They confirmed the fact that lissamine green disappears from the proximal tubule without reappearing in distal tubules, but found that when dilute trichloracetic acid was layered on the kidney surface after the disappearance of dye, the tubule fluid became green again. The seeming loss of lissamine green could be accounted for, either in part or in full, by decolorization of the dye in contact with poisoned tubular epithelium. There was, then, no firm evidence for lissamine green leakage. Steinhausen et al did find, however, that inulin injected directly into individual proximal tubules of mercury-poisoned rats appeared in ureteral urine from both kidneys. This could only occur if the injected inulin had leaked from the tubule into which it was injected, was recirculated, and then was filtered by both kidneys. Although this appeared to confirm Bank's (25) finding of inulin leakage, the validity of the experimental method is highly questionable. Mercury poisoning should affect both left and right kidneys equally; in this study there was a marked difference in function between the two kidneys of most animals studied. In five of the six rats, the urine output on the side prepared for micropuncture averaged only one third of that excreted by the contralateral kidney. Furthermore, inulin absorbed from the tubule into which it is injected should be excreted equally by the two kidneys after recirculation and filtration. In four of the six animals studied here, inulin excretion in ureteral urine of the injected kidney averaged only 25% of the total amount excreted. These results must, therefore, be taken with serious reservations. As pointed out by Biber (26) and Henry (28) and co-workers, solutions injected into friable, injured tubules might leak out artifactually while spontaneously filtered fluid does not. The possibility of artifactual leakage can be minimized by careful puncture with small, very sharp pipets and controlled injection under manometric control. Our own experience, despite these precautions, has led us to scrupulously avoid using tubular injections as an indicator of the permeability of necrotic tubules.

Barenberg and co-workers (29) measured single nephron GFR in rats during the first 5 hr after injection with 15 mg/kg BW $HgCl_2$. SNGFR was reduced one half within the first hour, and no correlation was found between the site of fluid collection and SNGFR to suggest leakage of the inulin indicator. The authors concluded that impaired filtration was responsible for renal failure in that model and that tubular factors were not involved. Since, however, the study was performed so soon after mercury injection, the authors recognized that this might reflect a transient hemodynamic effect of mercury and that other factors might be brought to bear at a later time.

Henry and co-workers (28), in studies of rats injected with 15 mg/kg BW potassium dichromate 24 hr earlier, found two populations of nephrons: one with sufficient flow for quantitation and another with tubular collapse or exceedingly poor flow. Single nephron GFR in the group of nephrons with the best function was only 21% of control. Proximal intratubular pressure of the entire nephron population was normal but certain tubules which apparently were obstructed had a mean pressure 8 mm Hg higher than control. Glomerular capillary pressure measured by the method of Gertz (30) was one fourth less than normal. Proximal tubule reabsorption half-time, an index of fractional water absorption, was prolonged almost sixfold. This finding, coupled with failure to observe tubular leakage of dyes injected i.v. or into the tubule, led the authors to conclude that passive backflow is not a prominent feature of dichromate-induced acute renal failure. Reduced filtration pressure was considered to be the primary determinant of the decrease in GFR. Although obstruction appeared to contribute to the production of oliguria in some nephrons, it was not a necessary feature of filtration failure. Chronically salt-loaded rats in that study failed to develop the histopathological alterations of tubular necrosis. Their surface nephrons did not show evidence of either obstruction or tubular collapse in vivo, and single nephron GFR values were almost four times greater than those of the best functioning nephrons of dichromate-poisoned animals that were not salt-loaded.

These authors also studied the renal function of rats in the developmental phase of globin-induced renal failure. Micropuncture was performed approximately 3 hr after globin was injected intravenously into rats on high or low salt intakes. Nephrons on the kidney surface rapidly became dilated and evidently were obstructed. Numerous "casts" were present in distal tubule segments on histologic examination. Mean whole kidney GFR of rats on both sodium regimes was approximately 15% of control, while SNGFR values were 61 and 46% of control in salt-deprived and salt-loaded rats, respectively. The disparity between whole kidney and single nephron GFR was attributed to an artifactual overestimation of SNGFR as a result of the process of collection from obstructed nephons. Unlike the dichromate-injected rats, globin-injected animals had modestly elevated proximal tubule hydrostatic pressures which, in the presence of decreased flow, is clear evidence of tubular obstruction. Proximal tubule transit time was approximately three times normal. Reabsorptive half-time, an indication of water absorption capacity, was prolonged in approximately one half of tubules studied and essentially normal in the remainder. No evidence of back leakage was observed when lissamine green was injected into tubules during intratubular measurement or when dye was injected intravenously for transit time measurements. The authors concluded that their finding of decreased reabsorptive capacity, glomerular filtration rate, and glomerular capillary pressure all indicated that the oliguria of these models of acute renal failure is most apt to be the result of renal cortical ischemia and a primary reduction in glomerular filtration. Although tubular obstruction was a prominent feature of rats injected with globin, it was by no means present in all nephrons and is not a necessary feature of filtration failure.

Biber and co-workers (26) performed micropuncture and nephron microdissection three days after injecting rats with 5 mg/kg $HgCl_2$ or 15 mg/kg potassium dichromate. In this study, whole kidney inulin clearances of a given animal correlated closely with the proportion of nephrons on the kidney surface in which free flow was demonstrable. ^{14}C-inulin injected into such functioning nephrons appeared in the urine, indicating that these tubules were neither obstructed nor grossly permeable to inulin. SNGFR values obtained in tubules with free flow were normal, regardless of the degree of depression of whole kidney GFR, again indicating that these tubules were impermeable to inulin, at least up to the site of collection. Fractional water absorption derived from the tubule fluid:plasma inulin concentration ratios in the first two thirds of proximal tubule length was essentially normal. Water absorption from the inaccessible distal portion of the proximal tubule presumably was impaired, however, since the inulin concentration ratio at the beginning of the distal convoluted tubules was markedly depressed. GFR values obtained in distal tubule collections were the same as in proximally derived samples; there was, then, no indication of inulin loss between the two collection sites. Potassium dichromate-injected animals had markedly decreased fractional water absorption throughout the proximal tubule, perhaps reflecting the more extensive tubular injury produced with dichromate. This reduction in water absorption was not the artifactual result of inulin leakage since absorption was similarly diminished when estimated by injecting droplets of isotonic saline into tubule segments and directly observing the rate of absorption. Unlike normal tubules, however, those of poisoned rats showed an increase in droplet absorption rate when intratubular hydrostatic pressure was deliberately increased. As an additional test for inulin leakage across necrotic epithelium of dichromate poisoned rats, ^{14}C-inulin was injected into proximal tubules and urine was collected from the injected and contralateral kidney. Inulin absorption from 14 nephrons averaged 7%, an amount which the authors considered the upper limit of possible loss of inulin since the process of injecting fluid into these necrotic and friable tubules could easily lead to undetected leakage entirely artifactually. They did, however, report that dye solutions injected into the most necrotic tubules of both dichromate- and $HgCl_2$-poisoned animals at times were seen to leak from the tubule lumen into the renal interstitium and capillaries. The same objection to injection studies applies here, however. Since injection pressure was not monitored, it is not clear that such leakage would occur spontaneously under free flow conditions.

Proximal tubule pressure was normal in the majority of proximal tubules studied regardless of the rate of tubule fluid flow. The tubules with slowed flow and normal intratubular pressure were considered to be obstructed. Apart from the leakage of dye from the most severely necrotic tubules on direct injection, passive backflow of filtrate was not demonstrable directly in this study and, indeed, judged from proximal tubule absorption of injected droplets, fluid transport was three times slower than control. The finding of normal pressure in obstructed tubules suggested to the authors, however, that filtrate might be absorbed across damaged epithelium at such a rapid rate that intratubular pressure would not rise despite outflow obstruction. If filtration in these obstructed nephrons indeed persisted but absorption were so rapid that hydrostatic pressure did not rise, injected droplets should have flowed rapidly along the tubule to replace that absorbed regardless of the presence of distal obstruction. The finding of normal pressure in fluid-filled tubules with scant flow could be interpreted as a suggestion of markedly decreased, rather than increased, tubular water absorption.

Flamenbaum and co-workers (31) have obtained results on rats poisoned with 4.7 mg/kg BW $HgCl_2$ 24 hr earlier which put the data of Bank and co-workers (see above) into a different perspective. Here, proximal tubule fluid flow in most nephrons was markedly reduced or absent. Despite the near absence of spontaneous flow, however, proximal tubule fluid could be collected at a normal rate and SNGFR values were minimally depressed. It will be recalled that comparably high SNGFR values were reported by Bank et al (25) when collections were obtained from the first one fifth of the proximal tubule. In this study, however, the finding of near normal GFR values in tubules with scant flow prior to insertion of a micropipet indicates that SNGFR had increased greatly as the result of venting these nephrons during collection. Such tubules evidently were obstructed, but proximal tubule pressures were not elevated. A similar phenomenon was described by Henry et al (28) in dichromate-poisoned tubules. When lissamine green-stained isotonic saline was injected under manometric control at pressures 5 to 15 cm saline above baseline into the lumen of these tubules in an attempt to wash out any obstructive material, the dye appeared in distal tubule segments immediately following a mass of white debris. The normal hydrostatic pressure found in frankly obstructed tubules indicated that glomerular capillary pressure was very low. Since the obstructive debris could be dislodged at a mean intratubular pressure only 8.6 cm H_2O greater than control pressure, it may be assumed that obstruction would not have developed if normal intraglomerular pressure had been maintained. After relief of obstruction by injection, proximal tubule flow became essentially normal as indicated by a near normal lissamine green transit time. In an effort to detect any alteration in tubular permeability which might be present, the tubular absorption of mannitol was compared with that of inulin. Proximal tubule inulin:mannitol concentration ratios of mercury-poisoned rats were no different from those of controls. Summing up, this study showed greatly reduced filtration in tubules which were obstructed but had low intratubular pressures. Venting the tubules evidently effected filtration pressure, which rose to a value sufficient to give near normal filtration rates. This appears to be a unique feature of low-dose mercury poisoning. No evidence for passive backflow could be adduced.

A companion study by DiBona and co-workers (32) offered the strongest indication that tubular permeability alterations contribute minimally to the pathogenesis of acute renal failure after the dose of mercury used by Flamenbaum (31), Bank (25), and Steinhausen (27). These authors performed studies in rats given 1% saline in place of tap water for one month to achieve maximal renin depletion. Whole kidney inulin and mannitol clearance of water-drinking rats both fell by 94%. Saline-loaded rats, by contrast, had an insulin clearance that was minimally different from control although their kidneys showed tubular necrosis which was just as severe as, but slightly less extensive than, that seen in their water-drinking counterparts. If the passive backflow theory were correct, the wide zone of frank tubular necrosis found in these saline-drinking animals should have provided ample opportunity for tubule fluid leakage. With near normal whole kidney inulin clearances, significant leakage of inulin could not have occurred even from these kidneys exhibiting frank necrosis, epithelial sloughing, and the denuding of basement membranes.

Oken and co-workers (33) have studied the development and established phases of glycerol-induced renal failure in normally hydrated and dehydrated rats injected intramuscularly with 10 ml/kg BW 50% glycerol. This maneuver which causes myolysis, acute plasma volume depletion, hemolysis, and myohemoglobinuria has many similarities to crush syndrome in man. Functional alterations were evident within 10 min after glycerol was injected. Stained oil droplets injected into individual tubules 2 to 4 hr after glycerol injection remained at the injection site, showing flow in the overwhelming majority of nephrons to be exceedingly slow or totally absent. GFR and fractional water absorption of the occasional nephron with sufficient flow to permit quantitation were both approximately 40% of normal, and fractional water absorption was not yet reduced. The intratubular pressure of fluid-filled nephrons was grossly subnormal. While pressure could not be measured in collapsed tubules, the fact that it was possible to inject fluid into them at pressures of 1 to 5 cm H_2O indicated that obstruction by external compression (e.g. edema) was not responsible for their collapse. Hemoglobin-stained fluid, present in most open tubules at this time, could usually be washed down the lumen by injecting isotonic saline at physiologic hydrostatic pressures.

Studies performed 19 to 24 hr after glycerol injection showed a smaller proportion of nephrons that were devoid of flow, but SNGFR and flow rates in the best functioning nephrons still were grossly abnormal. Proximal tubule hydrostatic pressure and fractional water absorption, where measurable, remained subnormal. Single nephron function in dehydrated rats was still too low to permit quantitation, but droplets of saline injected into lumens of collapsed tubules were so slowly absorbed as to be barely appreciable to the eye. In short, this study showed persistent failure of glomerular filtration in tubules whose pressure was low and fluid absorptive capacity was grossly reduced.

Ruiz-Guiñazù and co-workers (34) obtained results in rats with acute renal failure produced by intravenous injection of human methemoglobin that were very similar to those reported by Oken and co-workers (33). The animals were deprived of drinking water for 72 hr before methemoglobin was administered. Methemoglobin appeared in tubular lumens shortly after injection and proximal tubules rapidly became dilated, apparently as the result of increased resistance to tubule fluid outflow. On the following day, however, most proximal tubule segments on the kidney surface were essentially devoid of fluid. Flow in the few fluid-containing tubules usually was grossly reduced, and lissamine green solutions injected into collapsed tubules were very slowly absorbed; some tubules retained the dye solution for as long as 1 hr. Recovery of filtration apparently had begun within 48 hr in many nephrons. Proximal tubule hydrostatic pressure in the immediate postinjection period was some 8 mm Hg higher than control. Subsequently, intratubular pressure of dilated tubules remained elevated while the pressure of nondilated tubules, which were in the majority, was only approximately one half of control.

The hydrostatic pressure of "welling" vessels, an index of efferent arteriolar pressure, was measured directly at various times after methemoglobin injection. Average control pressure was 18.3 mm Hg, pressures measured 24 to 72 hr after injection being approximately one half of control. The finding of tubular collapse in most nephrons, low intratubular and efferent arteriole pressures, and slow peritubular capillary flow rates suggested to Ruiz-Guiñazù and co-workers that "the only single alteration that accounts for all our data is an increase in preglomerular resistance."

Jaenike (35) also performed studies on the methemoglobin-induced renal failure model, but used approximately one fourth the amount of methemoglobin after the rats were deprived of water for either 24 or 72 hr. As reported by Ruiz-Guiñazù (34), the vast majority of surface nephrons of the rats with more severe water deprivation were collapsed. Mean proximal hydrostatic pressures were significantly raised in most nephrons. Effective glomerular capillary pressure evaluated indirectly both by the tubule obstruction method of Gertz et al (30) and by simultaneous use of mannitol infusions and ureteral ligation was greatly reduced. Single nephron GFR values of the best functioning nephrons of the less dehydrated animals were less than one half of control. Fractional water absorption was normal, however, suggesting that there was no increase in outward diffusion of filtrate from these tubules. Single nephron GFR was too low for measurement in collapsed tubules, but previously collapsed tubules frequently became fluid-filled and developed supernormal pressures after mannitol was infused intravenously. The author concluded that tubule collapse was the result of filtration failure rather than supernormal absorption of filtrate since the small transtubular mannitol gradient obtained should have had little effect on fluid absorption, especially across pathologically permeable tubular epithelium. Reduced glomerular filtration pressure was considered the predominant factor in this model of acute renal failure, but tubular obstruction also was operative.

Wilson and co-workers studied the effect of a variety of infusions on the development of glycerol-induced, myohemoglobinuric acute renal failure in previously dehydrated rats (36). Plasma, mannitol, saline, and combined therapy with mannitol and saline at the time of glycerol challenge all decreased the severity of azotemia below that obtained in untreated rats. The BUN concentration of animals treated with any of these agents alone rose to approximately one half of the expected value. Animals given combined treatment with both mannitol and saline experienced a far more modest and transient BUN rise. On the day after glycerol injection, SNGFR values of all groups of rats were the same, regardless of their whole kidney GFR. The proportion of nephrons on the kidney surface with sufficient flow to permit quantitation varied greatly between groups, however, and correlated well with overall renal function. Only an occasional nephron of untreated glycerol-injected rats had flow that was sufficient for quantitation; at the other extreme, adequate flow was present in some 90% of nephrons on the kidney surface of animals given mannitol and saline simultaneously. Intratubular pressure was decidedly subnormal, indicating that diminished GFR was not the result of tubule obstruction. Fractional water absorption in the proximal tubules of functioning nephrons was significantly reduced in all groups of rats, including those given combined solutions.

In short, this study showed that the prophylactic effectiveness of plasma, saline, or mannitol infusions did not relate to tubular factors. Mannitol alone at a dose of 1.75 g/kg BW was no more protective than isotonic saline in terms of individual nephron GFR, fractional water absorption, and intratubular pressures. It seems doubtful, therefore, that the protection observed with mannitol was a result of an effect on urine volume or cell swelling alone. That combined therapy with saline and mannitol was so much more efficacious than either agent alone is best explained by the marked expansion of body fluid extravascular compartments and hemodynamic factors.

Dehydration is a potent factor predisposing to the development of most models of experimental acute renal failure as well as to the syndrome which occurs in man. This effect frequently has been attributed to increased urine concentration and slowed flow through distal segments of the nephron promoting tubule obstruction. The prevention of acute renal failure by maneuvers which will minimize the presence of pigment or cell debris had led to the theory that obstruction may cause reduced glomerular filtration as reviewed in this paper. The alternate possibility that potentially obstructive material may persist in tubule lumens as a result of filtration failure is exemplified in most micropuncture studies. Wilson and co-workers (37) have studied the role of concentrating mechanisms in the development of myohemoglobinuric renal failure in rats with complete hypothalamic diabetes insipidus. These rats, prior to glycerol injection, excreted volumes of urine equivalent to 80% of their body weight. If the prime effect of dehydration on acute renal failure were mediated through the renal concentrating mechanism, such rats, excreting massive volumes of dilute urine, should have been protected from renal insufficiency. They were not, and, indeed, showed more severe acute renal failure than normal rats subjected to glycerol. On micropuncture, renal functional abnormalities were identical to those obtained in normal rats at a comparable time after the same dose of glycerol. According to these results, it is apparent that the development of this model of acute renal failure is independent of the renal concentrating mechanism, and does not depend upon slowed flow through distal tubule segments with increased passive absorption of fluid, nor to the urine concentration per se. Factors other than the antidiuretic hormone titer, high urine output, and low urine flow rates must be the important mediators of the effect of dehydration in this model of acute renal failure.

Micropuncture studies performed in the developmental stage and the period of sustained renal insufficiency in glycerol-induced myohemoglobinuric acute renal failure have shown almost complete cessation of glomerular filtration (33), tubular collapse, low intratubular pressures, and no evidence of passive backflow filtrate. Since, however, most tubules on the kidney surface were collapsed and devoid of fluid, tubular integrity could not be properly assessed. In an attempt to characterize the events associated with recovery and to further evaluate the role of tubular factors which might relate to the characteristic slowness of the recovery phase, Oken and co-workers (38) have studied the recovery phase of glycerol-induced acute renal failure. Glycerol was administered in the same dose and route as in studies described earlier (33). The BUN concentration was measured serially and animals whose BUN concentration had fallen 20% or more below the peak value were considered to be in recovery and then were subject to micropuncture experiments. The degree of renal dysfunction ranged from a modest impairment to moderately severe azotemia. Two populations of nephrons were found: nephrons with perfectly normal GFR values and others with function which was too poor to permit quantitation. Each rat's BUN concentration at the time of study was inversely correlated with the proportion of functioning surface nephrons. Proximal tubule water absorption of the functioning nephrons was significantly depressed as manifest by the supernormal volumes of fluid obtained in timed collections and by low tubule fluid:plasma inulin concentration ratios. The mean proximal tubule hydrostatic pressure in fluid-filled nephrons was normal while that of collapsed tubules was decidedly low. There was, then, no indication that tubular obstruction significantly impaired the return of glomerular filtration. Since single nephron GFR values were normal and the volume of fluid collectable from the proximal tubule was actually supernormal, there was no evidence for a factitious loss of fluid as predicted by the passive backflow theory. It seemed unlikely that recovery might relate directly to the repair of tubule leakage or one would otherwise expect a wide spectrum of filtration and flow rates according to the degree of tubular repair. In view of the homogeneity of the values obtained, it is evident that such a spectrum of tubular permeability did not exist.

Interstitial edema has been considered a potential factor delaying recovery from acute renal failure (39). Renal edema is a universal concomitance of the recovery phase of glycerol-induced acute renal failure but in no way correlated with the completeness of return of renal function. In one animal, for example, a massively edematous kidney which weighed 1.51% of body weight was found in an animal whose BUN concentration had already fallen to near normal levels. It seems unlikely, therefore, that interstitial edema was a significant determinant to recovery in this study of glycerol-induced acute renal failure. Instead, recovery reflected the recruitment of function in increasing numbers of nephrons which had minimal function in the oliguric period. Recovery of individual nephron GFR, once begun, was rapid and complete. Fractional water absorption was grossly reduced, rather than enhanced. This was not due to urea-induced osmotic diuresis or expansion of extracellular fluid, occurring regardless of the animal's weight change and blood urea nitrogen concentration, and presumably reflected incomplete repair of tubular epithelium.

In conclusion, of the studies outlined in this review, several have directly examined single nephron GFR (24–26, 28, 29, 33–38), tubular pressures (24, 26, 28, 31, 33–38), fractional water absorption (24, 26, 28, 31, 33, 35–38), and tubule fluid flow rates (24, 26, 28, 31–38), and many have sought evidence of sufficiently increased tubular permeability to satisfy the passive backflow theory (24–29, 31–38). All studies but three (25, 26, 38) in which SNGFR has been measured quantitatively have shown glomerular filtration to be very severely depressed. The normal GFR values reported by Bank and co-workers were found by Flamenbaum et al (31) to be the artifactual result of the process of collection itself. The other studies in which near normal SNGFR was found, those of Biber (26) and Oken (38) and their co-workers, were done on animals already in the recovery state. In both studies, two populations of nephrons were found: those with normal SNGFR values and others with slight or virtually absent filtration. Recovery evidently is manifest by the progressive recruitment of increasing numbers of previously nonfiltering nephrons which rapidly attain a near normal GFR while the remainder filter minimal volumes. In short, there is little doubt that filtration virtually ceases during mercury and dichromate poisoning or methemoglobinemic and glycerol-induced myohemoglobinuric acute renal failure, and recovers rapidly with the onset of recovery.

Tubule obstruction of varying degrees of significance has been found in a number of studies (24, 26, 31, 33–35). Since, however, intratubular pressure has been decidedly low in most studies where SNGFR is maximally depressed, obstruction cannot be considered an essential factor in filtration failure. Indeed, the question is raised whether the lodgment of "obstructive" material is not more apt to be the result, rather than the cause, of impaired filtration. Nonetheless, at any given glomerular capillary pressure, intratubular pressure is an important determinant of filtration so that obstruction may, at least theoretically, contribute to impaired filtration in certain renal failure models.

Passive backflow has been incriminated in only 3 of the 15 micropuncture studies reported to date. In one (27), evident technological problems make the results difficult to interpret. In the second (25), results which could well have signified passive backflow were found by others (27, 31) to have a quite different explanation. In the third (26), the leakage of dyes injected into the most severely necrotic tubules was interpreted as an indication that such tubules would leak if filtration had been maintained. Whether that is so or not, the same authors found scant leakage of inulin injected into more representative nephrons, a grossly diminished rate of absorption of injected saline droplets, as well as equal SNGFR values in proximal and distal tubule samples collected from nephrons with low proximal tribule TF/P $_{IN}$ values. The possibility that passive backflow might have played a major role in the development of renal failure must, therefore, be considered in the light of the evidence against significantly increased fluid absorption in the same study.

The severe depression of glomerular filtrate found in so many studies must be considered as the essential feature of the models which form the basis of this review. Occurring with or without evidence of tubular obstruction, depressed filtration must be attributed to either a hemodynamic aberration or grossly altered glomerular capillary hydraulic conductivity. These possibilities are not mutually exclusive, but the well-documented decrease in renal blood flow associated with renal failure of this type suggests that low intraglomerular pressure is most apt to be responsible. Hence, we have used the term vasomotor nephropathy (40) in preference to the currently used names "acute renal failure" and "acute tubular necrosis."

ACKNOWLEDGMENT

It is a pleasure to acknowledge the secretarial assistance of Ms. Linda M. Don.

Literature Cited

1. Bancroft, F. W. 1925. *Ann. Surg.* 81:733
2. Frankenthal, L. 1916. *Virchows Arch.* 222:332
3. Minami, S. *Virchows Arch.* 1923. 245:247
4. Husfeldt, E., Bjering, T. 1937. *Acta Med. Scand.* 91:279
5. Yuile, C. L., Clark, W. F. 1941. *J. Exp. Med.* 74:197
6. Bywaters, E. G. L., Beall, D. 1941. *Brit. Med. J.* 1:427
7. Eggleton, M. G., Richardson, K. C., Schild, H. O., Winton, F. R. 1943. *Quart. J. Exp. Physiol.* 32:89
8. Hackradt. *Inaug.-Diss. München 1917.* Cited by Minami, S. 1923. *Virchows Arch.* 245:247
9. Trueta, J., Barclay, A. E., Daniel, P. M., Franklin, K. J., Prichard, M. M. L. 1947. *Studies of the renal circulation.* Springfield, Illinois: Thomas
10. Munck, O. 1958. *Renal Circulation in Acute Renal Failure,* 22, 35. Oxford: Blackwell
11. Chedru, M.-F., Baethke, R., Oken, D. E. 1972. *Kidney Int.* 1:232
12. Hollenberg, N. K., Adams, D. F., Oken, D. E., Abrams, H. L., Merrill, J. P. 1970. *New Engl. J. Med.* 282:1329
13. Marshall, D., Hoffman, W. S. 1949. *J. Lab. Clin. Med.* 34:31
14. Sirota, J. H. 1949. *J. Clin. Invest.* 28:1412
15. Selkurt, E. E. 1945. *Am. J. Physiol.* 144:395
16. Phillips, R. A., Hamilton, P. B. 1948. *Am. J. Physiol.* 152:523
17. Richards, A. N., Westfall, B. B., Bott, P. A. 1936. *J. Biol. Chem.* 116:749
18. Hayman, J. M., Richards, A. N. 1926. *Am. J. Physiol.* 79:149
19. Bobey, M. E., Longley, L. P., Dickes, R., Price, J. W., Hayman, J. M. Jr. 1943. *Am. J. Physiol.* 139:155
20. Laake, H. 1945. *Acta Med. Scand. Suppl.* 168
21. Oken, D. E. *Am. J. Med.* In press
22. Schmidt, U. et al 1972. *Proc. 5th Int. Congr. Nephrol. 1972,* p. 165
23. Tanner, G. A., Sloan, K. L., Sophasan, S. 1973. *Kidney Int.* 4:377
24. Flanigan, W. J., Oken, D. E. 1965. *J. Clin. Invest.* 44:449
25. Bank, N., Mutz, B. F., Aynedjian, H. S. 1967. *J. Clin. Invest.* 46:695
26. Biber, T. U. L. et al 1968. *Am. J. Med.* 44:664
27. Steinhausen, M., Eisenbach, G. M., Helmstädter, V. 1969. *Pflügers Arch. Gesamte Physiol.* 311:1
28. Henry, L. N., Lane, C. E., Kashgarian, M. 1968. *Lab. Invest.* 19:309
29. Barenberg, R. L., Solomon, S., Papper, S., Anderson, R. 1968. *J. Lab. Clin. Med.* 42:473
30. Gertz, K. H., Mangos, J. A., Braun, G., Pagel, H.D. 1966. *Pflügers Arch. Gesamte Physiol.* 288:369
31. Flamenbaum, W., McDonald, F. D., DiBona, G. F., Oken, D. E. 1971. *Nephron* 8:221
32. DiBona, G. F., McDonald, F. D., Flamenbaum, W., Dammin, G. J., Oken, D. E. 1971. *Nephron* 8:205
33. Oken, D. E., Arce, M. L., Wilson, D. R. 1966. *J. Clin. Invest.* 45:724
34. Ruiz-Guiñazù, A., Coelho, J. B., Paz, R. A. 1967. *Nephron* 4:257
35. Jaenike, J. R. 1969. *J. Lab. Clin. Med.* 73:459
36. Wilson, D. R., Thiel, G., Arce, M. L., Oken, D. E. 1967. *Nephron* 4:337
37. Wilson, D. R., Thiel, G., Arce, M. L., Oken, D. E. 1969. *Nephron* 6:128
38. Oken, D. E., DiBona, G. F., McDonald, F. D. 1970. *J. Clin. Invest.* 49:730
39. Merrill, J. P. 1960. *Ann. Rev. Med.* 11:127
40. Oken, D. E. 1971. *Nephron* 8:505

Copyright 1975. All rights reserved

THE STEAL SYNDROMES[1]

❖7153

James F. Toole, M.D. and Charles Patrick McGraw, Ph.D.
Department of Neurology, Bowman Gray School of Medicine of Wake Forest University, Winston-Salem, North Carolina 27103

INTRODUCTION

"... it is the blood in quantity rather than blood under pressure that most organs require." Jarisch, cited by Franklin (1).

In 1960, Contorni (2) reported that a stenotic lesion in the subclavian artery proximal to the origin of the vertebral artery could result in reversed vertebral artery flow. In 1961, vertebrobasilar insufficiency resulting from this reversal of flow was documented for the first time (3) and the syndrome was labeled "the subclavian steal syndrome" (4). Because this form of cerebrovascular insufficiency was correctable, and because once discovered it was recognized frequently, numerous publications followed (5–9), with colorful names such as "vertebral grand larceny" (10), "bilateral steals" (11, 12), "double steal" (13) and "triple steal" (14), "intracranial steal" (15), "the Robinhood syndrome" (16), "occipital vertebral steal" (17), "the ophthalmic steal" (13, 17), and "asymptomatic steal" (13). Furthermore, the siphoning of blood from its destined area of distribution by a reversal of flow was detected in other vascular beds such as the aorto-iliac (18, 19), celiac (20), hepatic (20, 21), mesenteric (19, 20, 22, 23), spinal (24), coronary (19, 25, 26), renal splanchnic (27), and radial-ulnar anastomotic beds (28).

PATHOPHYSIOLOGY

Steals occur when the distal segment of an obstructed artery is perfused through collaterals that divert blood from adjacent arterial territories. They develop whenever there are (*a*) gradients from high to low pressure systems and (*b*) adequate arterial or arteriolar anastomoses. These anastomoses become dilated for several reasons, the most apparent being the development of ischemia from the accumulation of end products of metabolism such as CO_2. In some instances, the development of ischemic symptoms in the donor distribution leads to false localization of the disease.

---

[1] Supported in part by U.S. Public Health Service Grant NS 06655 and the Neurosurgical Research and Development Fund.

As mentioned above, steals are not limited to a particular vascular distribution; they can also occur with or without symptoms. As a consequence, it has been suggested (19) that the term "steal syndrome" be used only when diversion of flow is accompanied by symptoms of deficient circulation and that "steal phenomenon" or "steal effect" be used when the steal is asymptomatic.

The extent of collateral pathways is sufficient to explain asymptomatic steals. For example, angiography has shown (6) that the following channels divert blood to the distal subclavian artery when the proximal subclavian or innominate artery is occluded: (*a*) the contralateral vertebral artery at the vertebrobasilar junction; (*b*) either internal carotid artery through the posterior communicating arteries via the basilar artery; (*c*) the occipital branch of the ipsilateral external carotid artery and the muscular branches of the vertebral artery; (*d*) the contralateral vertebral artery through anastomoses with the muscular branches of the two vertebral branches; (*e*) anastomoses between muscular branches of the vertebral artery and the branches of the costocervical or thyrocervical trunks of the thyroid axis on the two sides; (*f*) the internal mammary and intercostal arteries; and (*g*) the thyroid arteries.

In this chapter we will consider mainly cerebral steal syndromes, that is, steals that include any diversion of blood from an intracranial structure to any other structure, that diversion resulting in a loss of adequate cerebral perfusion and, therefore, adequate cellular exchange.

INTRACEREBRAL STEALS

Cerebral steals become symptomatic only when enough blood is siphoned away from the brain to produce ischemia. Many collateral channels such as those enumerated above can be utilized to spare the brain. However, it is possible that the so-called asymptomatic steals are in reality causing ischemia, but are doing so in an area that does not produce recognized symptoms. One such location could be the parietal association areas of the brain.

Even though steal is present constantly, it may become symptomatic intermittently because the needs of tissues fluctuate. Symptoms of vascular insufficiency occur only if the metabolic needs of the tissue exceed the amount of available blood.

Sometimes, dilation in a local cerebrovascular bed results in the shunting of blood from ischemic areas into normal zones (15). For example, normal arterioles might be dilated by excessive CO_2 while arterioles in damaged tissue might not be so affected. Flow would then be diverted into the normal tissue and might precipitate cerebrovascular insufficiency in the marginally perfused regions. This increased flow in the zone of brain surrounding an infarcted area has been called "luxury perfusion" and that in the infarction, "poverty perfusion" (28). In normal brain, autoregulation compensates for such changes in perfusion pressure and the cerebral blood flow remains constant. However, in ischemic areas, vessels lose their autoregulative capacity and blood flow varies passively with arterial pressure.

Most investigators agree (16, 29–31) that blood which normally supplies the ischemic area is diverted by the great reduction in cerebrovascular resistance induced by the dilated arteries in the presence of CO_2. Lassen and co-workers (16, 29) have shown that an area of ischemic brain with a lowered regional blood flow but normal arterial PCO_2 may have a further reduction of flow in response to increased arterial PCO_2 rather than an elevation as one might expect.

This reduction has been observed in abnormal areas of the brain surrounding arteriovenous malformations and in brain neoplasms (30, 32). Samuel, Grange & Hawkins (33) reduced arterial PCO_2 and produced vasoconstriction of surrounding normal brain without concomitant constriction of tumor vessels. The result was relative hyperemia of the tumor and ischemia of normal brain. Such steals may explain unexpected clinical signs in tumor patients.

In addition to a decrease in blood flow to an ischemic zone in response to an increase in arterial PCO_2, there is a decrease in perfusion pressure as well (34, 35), probably because the vascular bed is already maximally dilated and is therefore incapable of responding further to changes in arterial PCO_2 or PO_2. This conclusion is supported by the finding of Kindt & Youmans (36) that, with progressive stenosis of neck vessels, the response of flow in normal vessels to inhalation of CO_2 gradually declines, a fact that suggests that the dilating effects of occlusion and hypercapnia affect the same mechanism. The extent to which this reserve vasodilator capacity is used up by occlusion or hypercapnia depends upon their relative degree of each (31). Vasodilator capacity is reduced by elevation of intracranial pressure: the higher the baseline intracranial pressure, the greater the further elevation of intracranial pressure in response to an elevation in arterial PCO_2, because the intracranial adaptive capacity is already used up by whatever is elevating the intracranial pressure. During the initial postocclusion period, there is reduced total cerebral flow and very low flow in the infarcted zone, loss of local autoregulation, diminished reactivity to CO_2, and increased tissue pressure and intracranial pressure due to edema. During this period, vasodilator therapy may result in a steal with spreading of the zone of infarction (32). After two weeks, however, CO_2 reactivity may return, edema will subside, and vasodilators may then have a therapeutic effect.

Another phenomenon that might be explained on the basis of abnormal responsiveness to CO_2 is the headache secondary to elevated intracranial pressure that occurs in patients with brain tumors. This headache occurs predominantly during the night when the arterial PCO_2 is slightly increased. Symon (30) has suggested that it may be due not only to an increase in intracranial pressure from the elevated PCO_2 but also to a diversion of blood away from an area of abnormal circulation in the tumor, with perhaps a secondary swelling of the tumor due to ischemia. Under normal circumstances, an intrinsic autoregulatory capacity would readily compensate for the increase in PCO_2, but an intracranial mass partially uses up this adaptive capacity and distorts the effects of CO_2.

Even slight CO_2 retention, when associated with pulmonary disease, secondary hypercapnia, or sleep (37), plus diminished autoregulatory reserve in hypertensive patients, sets off brain swelling and progressive elevation of intracranial pressure. Secondary pressure on brain stem respiratory centers further depresses respiration, and arterial PCO_2 increases even more. Secondary ischemia and further brain swelling occur until there is enough cerebral ischemia to activate autonomic reflexive catecholamine release, systemic vasoconstriction, and a rise in arterial blood pressure with adequate cerebral perfusion. A similar series of events could be used to explain transient increases in intracranial pressure such as plateau waves (38). Respiratory irregularity or increase in percent expired CO_2 causes such slight dilation in a patient whose vasodilator capacity has been used up that such increases could set off a progressive elevation of intracranial pressure, a diversion of blood away from the damaged area, and secondary ischemia and swelling that progress until they trigger a release of catecholamines (39), cerebrovascular constriction, and termination of the plateau wave heralded by an initial increase in heart rate (40). This mechanism has been called the "intracranial squeeze" (41, 42). That such vasodilation occurs during plateau waves is supported by the finding of an increase in intracranial blood volume during such waves (43).

In 1968, Lassen & Palvölgyi (16) described a "reverse intracerebral steal" in which hypocapnia increased cerebral blood flow in ischemic areas. Because blood was diverted from the "rich" areas and given to the "poor" areas, they called the result the Robinhood syndrome. Blood flow was increased in the nonreactive dilated ischemic vasculature by constriction of normal vessels. They believed that the inverse steal might also have been due to a reduction in intracranial pressure with secondary increase in perfusion. Other authors (44) disagree that arterial PCO_2 increases when ischemia is present because blood flow either does not change or decreases in an ischemic region. It is probably best that hypercapnia not be induced in the patient with cerebral infarction because intracerebral steal, luxury perfusion, or inversed steal may occur unless the intracranial pressure is normal.

SUBCLAVIAN STEALS

The best known steals are those that were, at first, believed to be caused only by subclavian obstruction. Actually, reversed flow in the vertebral artery can be caused by obstruction in the proximal subclavian artery, the brachiocephalic artery, or the aortic arch; reversed flow in the carotid artery can be caused by brachiocephalic obstruction; and flow can be diverted from the vertebrobasilar to the carotid system. However, "subclavian steal syndrome" with the site of obstruction mentioned is the term most commonly used. The direction in which the blood flows when the various obstructive lesions exist depends upon anastomoses and pressure gradients. Blood pressure is normally higher at the origin of the carotid and vertebral arteries than it is in their more distal branches and, consequently, blood flows toward the brain. However, when the pressure in the proximal artery is reduced by 10%, blood flow in the distal branches decreases, and when the reduction exceeds 20%, flow reverses direction. In the case of vertebral artery flow, this reduction in pressure and reversal of flow is increased due to dilatation of the muscular bed when the arm is exercised. This feature, a pulse delay, a subclavian-vertebral bruit, and a difference in blood pressure of at least 20 mm Hg between the two arms are the four components of the subclavian steal (45).

The average age of patients with subclavian steal is 50–60 years (46, 47) and there is a predominance in men (70%). The subclavian steal also occurs three times more often on the left side (47). Symptoms generally occur in older patients, seldom in children. Bilateral subclavian steal syndrome occurs in 1 to 2% of all patients. About 77% of all patients suffer from occlusion and about 23% from stenosis. Of the total, about one half suffer from multiple vascular occlusions or stenoses (46, 47) and 10% are hypertensive. Spontaneous improvement may occur in many cases (8), apparently due to the development of adequate collaterals.

Etiology

Atherosclerosis is the most common cause of subclavian steal syndrome. However, they can be the result of trauma, thrombosis, inflammation (48), and congenital anomalies (49). Steals have also been observed after a Blalock-Taussig procedure for tetralogy of Fallot, with congenital atresia of the subclavian artery, after ligation of the subclavian artery during resection of a traumatic thoracic aortic aneurysm, and after kinking of the subclavian artery associated with pseudocoarctation (50–59). McGuire & Tucker reported a subclavian steal caused by peripheral dilatation in a paretic arm (60).

A subclavian steal may reduce total cephalic blood flow by as much as 41%, which in normal brain is compensated for by cerebral autoregulation (61). However, in the presence of atherosclerosis or fixed hypertension, the vascular bed loses its autoregulatory capacity and vascular insufficiency develops (6). Therefore, symptomatic steals are the result of a combination of factors, including systemic vascular disease resulting in a lack of collateral channels, loss of compensatory cerebral autoregulation, and fluctuating demands for blood in the competing vascular beds. For example, cutaneous vasodilation and muscular exercise cause dilation of limb vascular beds and augment the siphoning effect. Symptoms will occur when the metabolic needs of the tissue exceed the amount of available blood. If the patient is to remain asymptomatic he must have a normal circle of Willis and normal extracranial vessels (60): an incomplete circle of Willis is found in a significantly higher proportion of patients with vertebrobasilar insufficiency than in normal patients.

Symptoms

The arm symptoms are due to ischemia and usually (with a frequency of about 40%) consist of paresthesias, feelings of coldness, and weakness. Rapid weariness during arm exercise, pain while resting, and skin pallor occur at about 20% frequency, but pain during exercise seldom occurs (8%). Fingertip gangrene is rare, but has been described, as have forearm and hand atrophy and petechial bleeding of the fingers (46). Javid et al (62) described fingertip necrosis in three of four patients with right-sided subclavian steals, but in none of 21 patients with left-sided subclavian steals.

The neurologic manifestations of a subclavian steal are usually those of vertebrobasilar insufficiency: dizziness (75%), visual defects (36%), syncope (25%), ataxia (11%), and heterogenous motor and sensory symptoms (11%). However, they may instead be symptoms referrable to total cerebral blood flow. Such symptoms may include headaches (44%), alterations in speech (11%), memory changes (8%), personality disorders, and insomnia. Cerebral symptoms alone are more frequent (47%) than cerebral and arm symptoms together (42%), which in turn are more frequent than arm symptoms alone (11%) (46, 47). The first symptoms often appear when the patient lifts or carries a heavy object, has a continuous load on the affected arm, or makes sudden arm or head movements. Isotonic contraction produces symptoms more frequently than do isometric contractions. The symptoms gain importance with recurrence, but an average of two years may pass between the first manifestation and diagnosis. The symptoms are not specific and may also be found with other cerebral vessel diseases (46).

Cameron & Wright (63) described a patient with subclavian steal who had olfactory hallucinations and vivid dream states that suggested temporal lobe dysfunction. These symptoms were relieved when the subclavian steal was corrected. Apparently, in that patient, the temporal lobe was the critical perfused area.

The most common complex manifesting subclavian steal comprises vertigo, dizziness, syncope, occipital headaches, arm pain, tinnitus, lightheadedness, blurred vision, ataxia, and unsteadiness.

The vertebral artery on the side opposite the stealing artery is the larger in most cases, due to compensatory arterial dilation in response to chronic increase in demand (64). Also, increases in compensatory flow of 78% in the vertebral artery and 43–45% in each carotid artery have been shown experimentally (3). Therefore, vascular insufficiency, but not permanent neurologic deficits, is the hallmark in pure steals. The existence of a permanent neurologic deficit suggests multiple vascular involvement, a suggestion supported by the high incidence of hypertension, arteriosclerotic heart disease, and peripheral vascular disease that is seen in conjunction with subclavian steals (65–67). Such neurologic deficits, when present, rarely remit spontaneously (52).

The radial pulse on the affected side is either delayed, diminished, or absent in 98% of patients (67). In eight patients with quantitated abnormal pulse delays ranging from 0.014 to 0.2 sec, all had complete occlusions, while the pulse delay in normal controls was 0.00 to 0.01 sec (68, 69). A delayed pulse arrival time has even been considered almost pathognomonic of the subclavian steal syndrome (45).

In one series of ten patients (70), the pulse volume of the ophthalmic artery was diminished on the side with reversed vertebral flow. There was a pressure drop in the ophthalmic artery and the severity of that drop appeared to depend on the degree of the reversed flow. These findings were aggravated when the affected arm was exercised. Such ophthalmic findings may be useful when one must assess the degree of functional cerebrovascular insufficiency (70).

The presence of a supraclavicular systolic bruit can be a helpful clue to the diagnosis of subclavian artery obstruction, since 60% of patients with such obstruction have these bruits. Because some normal patients have such bruits (67), a supraclavicular systolic bruit, by itself, has little diagnostic significance. The usual bruit is of moderate intensity and is systolic in time; some patients have an associated thrill. The bruit is usually well localized over the path of the subclavian artery. Neck rotation, bracing back the shoulders, and deep inspiration may change its characteristics and intensity by compressing the artery. Exercising of the affected arm can increase the intensity of the bruit or add a diastolic component.

A difference in brachial blood pressure between the two sides is also a necessary characteristic of unilateral subclavian steal. The difference in systolic readings may vary from 20 to 130 mm Hg but is most often between 30 and 60 mm Hg. The difference in diastolic readings can be much greater. Although a significant difference in brachial pressures may alert one to the possibility of a subclavian steal, this difference has no correlation with the degree of stenosis. Rahman & Rodbard (71) have concluded, after recording the time of appearance of the Korotkoff sounds in each beat and the intensity and duration of these sounds, that the severity of the obstruction can be estimated by that method. If a steal is bilateral, the blood pressure in the two arms is equally reduced compared with that in the legs (67). A technique for bilateral simultaneous sphygmomanometry has been described (72).

Arm exercise can precipitate or aggravate cerebral symptoms, depending upon: (*a*) severity and duration of arm exercise, (*b*) amount of vasodilation in the arm, and (*c*) adequacy of the arm collaterals. Such exercise may even precipitate a loss of pulse, a bruit, or vertebral basilar symptoms. Exercise of both arms is a more provocative test, because blood may be drawn simultaneously from both vertebrals and the amount of blood diverted will be greater. Repetitive arm exercise may also be more effective than sustained contraction because the pumping activity diverts greater quantities of blood.

In some cases ischemia of pectoral muscles produces angina-like symptoms. Perhaps this is due to reversal of mammary artery and thyrocervical flow.

Treatment

Surgery is the definitive therapy in patients who have recurring, disabling symptoms. However, the diagnosis is a clinical one and arteriography is only required to define the exact site and extent of the obstructive lesion prior to surgery. Theoretically, the ideal treatment is restoration of normal anatomy by thromboendarterectomy with, if necessary, patch grafting or replacement. However, some surgeons prefer to bypass the obstruction by grafting the subclavian artery to the aorta or the carotid artery. A carotid-subclavian anastomosis is advantageous because it can be performed through the supraclavicular approach without the need for a thoracotomy (45). Ligation of the vertebral artery, the subclavian artery, or both has been used to reduce the drain on the cerebral circulation and to shunt the blood into the brain. Such ligations can be done through a cervical approach in poor-risk patients under local anesthesia. Operations upon the carotid and brachiocephalic arteries should be most effective when the circle of Willis is intact. If the circle of Willis is interrupted, vertebrobasilar insufficiency will probably not be relieved by operations that augment carotid inflow. In these circumstances, correction of the steal itself or improvement of vertebral inflow on either side is more appropriate (64). An intrathoracic surgical bypass procedure that allows histologic examination of the subclavian artery is recommended when the cause of the syndrome is in doubt (48).

Resnicoff, DeWeese & Rob (73) have concluded that when the right subclavian artery is involved, endarterectomy through a cervical approach is the procedure of choice; left subclavian artery involvement is best handled by employing a carotid subclavian bypass; and vertebral ligation is best for either side in the poor-risk patient, because it can be done under local anesthesia. Others claim that a relatively simple bypass procedure can be done even in poor-risk patients (66). Arteriography should be done before an operation in which ligation of a vertebral artery is contemplated in order to indicate whether the contralateral vertebral artery is adequate to maintain the posterior cerebral circulation (3). An additional feature that may be helpful in diagnosis and that may indicate when a favorable surgical result can be expected is inflation to systemic pressure of a blood pressure cuff on the side of the steal. If such inflation alleviates the cerebral ischemic complaints, an operation should be successful (50). Concurrent stenosis of an internal carotid artery should be repaired before the subclavian steal is corrected. Reconstructive surgery for asymptomatic patients is not advocated.

SUMMARY

Blood can be stolen from almost any vascular bed and redistributed. The external carotid artery can steal from cerebral structures when occlusion of the common carotid artery proximal to the bifurcation causes reversal of flow through the internal carotid artery (13, 45, 46). The external carotid artery has also been shown to steal blood flow from the vertebral artery distribution (13, 74). Double steals involving both the subclavian and external carotid arteries are seen in 1 to 2% of patients (13, 46). There is also the "triple steal syndrome" wherein both subclavian arteries and one external carotid artery are involved (14).

Literature Cited

1. Franklin, K. J. 1949. *Proc. Royal Soc. Med.* 42:721–30
2. Contorni, L. 1960. *Minerva Chir.* 15:268–71
3. Reivich, M., Holling, H. E., Roberts, B., Toole, J. F. 1961. *N. Engl. J. Med.* 265:878–85
4. Editorial. 1961. *N. Engl. J. Med.* 265:912–13
5. Mannick, J. A., Suter, C. G., Hume, D. M. 1962. *J. Am. Med. Assoc.* 182:254–58
6. Patel, A., Toole, J. F. 1965. *Medicine* 44:289–303
7. Ramirez-Boettner, C. M., Sosa, O. D., Balanza, J., Calcena, R. M. 1966. *Ann. Intern. Med.* 64:142–44
8. Heyman, A. et al 1964. *Arch. Neurol.* 10:581–89
9. Irvine, W. T., Luck, R. J., Jacobey, J. A. 1965. *Lancet* 1:994–96
10. Daves, M. L., Treger, A. 1964. *Circulation* 29:911–13
11. Agee, O. F. 1966. *Am. J. Roentgenol Radium Ther. Nucl. Med.* 97:447–51
12. Coder, D. M., Frye, R. L., Bernatz, P. E., Sheps, S. G. 1965. *Mayo Clin. Proc.* 40:473–76
13. Barnett, H. J. M., Wortzman, G., Gladstone, R. M., Lougheed, W. M. 1970. *Neurology* 20:1–14
14. Cala, L. A., Armstrong, B. K. 1972. *Aust. N. Z. J. Med.* 3:275–77
15. Dyken, M. L. 1971–72. *Eur. Neurol.* 6:301–5
16. Lassen, N. A., Palvölgyi, R. 1968. *Scand. J. Lab. Clin. Invest.* 22:Suppl. 102, XIII:D
17. Huckman, M. S., Haas, J. 1972. *Am. J. Ophthalmol.* 74:1094–99
18. Kountz, S. L., Laub, D. R., Connolly, J. E. 1966. *Arch. Surg.* 92:490–97
19. Vollmar, J., von 1971. *Münch. Med. Wochenschr.* 14:501–6
20. Reuter, S. R., Olin, T. 1965. *Radiology* 85:617–27
21. Gryboski, J. D., Clemett, A. 1967. *Pediatrics* 39:344–47
22. Bücheler, E., Düx, A., Rohr, H. 1967. *Fortschr. Geb. Roentgenstr. Nuklearmed.* 106:313
23. May, R., Nissl, R. 1965. *Fortschr. Roentgenstr.* 103:743–44

24. Doppman, J. L., Di Chiro, G., Ommaya, A. K. 1969. *Selective Arteriography of the Spinal Cord.* St. Louis, Mo: Warren H. Green. 157 pp.
25. Baker, L. D., Pomposiello, J. C., Sharma, G. V. R. K., Snow, J. A., Messer, J. V. 1969. *Clin. Res.* 17:228 (Abstr.)
26. Editorial. 1970. *Circulation* 42:193–94
27. Alfidi, R. J. et al 1972. *Radiology* 102:545–49
28. Bussell, J. A., Abbott, J. A., Lim, R. C. 1971. *Ann. Intern. Med.* 75:387–94
29. Hoedt-Rasmussen, K. et al 1967. *Arch. Neurol.* 17:271–81
30. Symon, L. 1968. *Scand. J. Lab. Clin. Invest.* 22:Suppl. 102, XIII:A
31. Symon, L. 1969. *Int. Anesth. Clin.* 7:597–615
32. Wullenweber, R. 1968. *Scand. J. Lab. Clin. Invest.* 22:Suppl. 102:XIII:C
33. Samuel, J. R., Grange, R. A., Hawkins, T. D. 1968. *Anaesthesia* 23:543–51
34. Brawley, B. W., Strandness, D. E., Kelly, W. A. 1967. *Arch. Neurol.* 17:180–87
35. Symon, L. 1963. *J. Physiol.* 165:62P
36. Kindt, G., Youmans, J. 1968. *Scand. J. Clin. Lab. Invest.* 22:Suppl. 102, XVI:H
37. Cooper, R., Hulme, A. 1966. *J. Neurol. Neurosurg. Psychiat.* 29:564–70
38. Lundberg, N. 1960. *Acta Psychiat. Scand. Suppl.* 149:1–193
39. Wurtman, R. J., Zervas, N. T. 1974. *J. Neurosurg.* 40:34–36
40. Tindall, G. T., McGraw, C. P., Vanderveer, R. W., Iwata, K. 1972. *Intracranial Pressure,* ed. M. Brock, H. Dietz, 227–31. Heidelberg:Springer
41. Welch, K. M. A., Meyer, J. S. 1970. *Lancet* 1:1316
42. Editorial. 1972. *Med. J. Aust.* 2: 1442–43
43. Risberg, J., Lundberg, N., Ingvar, D. H. 1969. *J. Neurosurg.* 31:303–10
44. Yamaguchi, T., Regli, F., Waltz, A. G. 1971. *Stroke* 2:139–47
45. Toole, J. F., Patel, A. N. 1967. *Cerebrovascular Disorders,* 111–15. New York: McGraw-Hill
46. Heidrick, H., Bayer, O. 1969. *Angiology* 20:406–13
47. Santschi, D. R., Frahm, C. J., Pascale, L. R., Dumanian, A. V. 1966. *J. Thorac. Cardiov. Surg.* 51:103–12
48. Pollock, M., Blennerhassett, J. B., Clarke, A. M. 1973. *Neurology* 23: 653–57
49. Love, J. W. 1968. *J. Cardiovasc. Surg.* 9:358–64
50. Piccone, V. A., Karvounis, P., LeVeen, H. H. 1970. *Angiology* 21:240–59
51. Craddock, D. R., Logan, A. 1966. *Brit. J. Surg.* 53:926–29
52. Mandelbaum, I., Nahrwold, D. L., Dzenitis, A. J. 1967. *Ann. Surg.* 165:314–17
53. McClenathan, J. E., Voss, M. W., Brown, R. B., Long, D. M. 1963. *N. Engl. J. Med.* 268:517–20
54. Rojas, R. H., Levitsky, S., Stansel, H. C. 1966. *J. Thorac. Cardiov. Surg.* 51:113–15
55. Gerber, N. 1967. *Am. J. Dis. Child.* 113:709–13
56. Folger, G. M., Shah, K. D. 1965. *Circulation* 31:241–48
57. Killen, D. A., Foster, J. H., Gobbel, W. G. Jr., Scott, H. W. Jr. 1967: *Am. Surg.* 33:128–38
58. Lochaya, S., Kaplan, B., Shaffer, A. 1967. *Am. Heart J.* 73:369–74
59. Mozes, M., Bank, H. 1967. *Lancet* 2:533–34
60. McGuire, T. H., Tucker, E. J. 1971. *Int. Surg.* 56:72–77
61. Rapela, C. E., Bush, C. B., Franklin, T. D., Toole, J. F., Green, H. D. 1964. *Physiologist* 7:233 (Abstr.)
62. Javid, H., Julian, O. C., Dye, W. S., Hunter, J. A. 1965. *Arch. Surg.* 90:634–43
63. Cameron, D. J., Wright, I. S. 1964. *Ann. Intern. Med.* 61:128–33
64. Lord, R. S. A., Adar, R., Stein, R. L. 1969. *Circulation* 40:871–78
65. Killen, D. A. et al 1966. *J. Thorac. Cardiov. Surg.* 51:539–60
66. Crawford, E. S., DeBakey, M. E., Morris, G. C. Jr., Howell, J. F. 1969. *Surgery* 65:17–31
67. Rennie, L., Ejrup, B., McDowell, F. 1964. *Neurology* 14:751–56
68. Grossman, B. L., Brisman, R., Wood, E. H. 1970. *Radiology* 94:1–6
69. Conrad, M. C., Toole, J. F., Janeway, R. 1965. *Circulation* 32:346–51
70. Follmann, P., Littmann, J., Urai, L., Heszberger, S. 1969. *Opthalmologica* 1959:363–72
71. Rahman, A., Rodbard, S. 1970. *Arch. Intern. Med.* 125:1027–29
72. Toole, J. F., Tulloch, E. F. 1966. *Circulation* 33:952–57
73. Resnicoff, S. A., DeWeese, J. A., Rob, C. G. 1970. *Circulation* 41–42(Suppl. II):147–51
74. Peabody, C. N., O'Brien, B. 1966. *Angiology* 17:148–52

Copyright 1975. All rights reserved

THE ANTIGLOBULIN TEST IN AUTOIMMUNE HEMOLYTIC ANEMIA

❖7154

Wendell F. Rosse, M.D.
Hematology Division, Department of Medicine, Duke University Medical Center and the Durham Veterans Administration Hospital, Durham, North Carolina 27710

Autoimmune hemolytic anemia (AIHA) is a process wherein red cells are destroyed because of their interaction with an antibody made by the immune system of the patient himself. In this sense, immunological tolerance has been broken since the antigens against which the antibody is directed are those of the organism making the antibody (1). In a few instances these antigens have been either chemically or serologically defined, but for the most part they remain unknown.

The antibodies involved in autoimmune hemolytic anemia are usually either IgG or IgM (2). The distribution of the subclasses among the IgG antibodies does not correspond to the distribution of these subclasses in the IgG of serum. Antibodies of the IgG_1 and IgG_3 subclasses seem to predominate although IgG_2 and IgG_4 antibodies have been identified (3, 4). IgA antibodies have been identified (5, 6) but to date no IgD or IgE antibodies have been implicated in the causation of autoimmune hemolytic anemia.

In almost all instances the interaction between antigen and antibody results in the fixation to the membrane of immunoproteins, either antibody or the components of complement (7). The immunoglobulins may in some instances be detected by their ability to effect agglutination; this is usually not true of the immunoglobulins of the IgG class (7).

Antibody and the components of complement fixed to the red cell membrane may be detected by the antiglobulin reaction first proposed by Coombs (8). In this reaction antibodies produced in other animals (hetero-antibodies) against human immunoproteins are reacted with washed red cells. If immunoprotein molecules are fixed to the red cell membrane, agglutination will occur. This reaction may be used to detect free antibody in the serum if normal red cells are incubated with the serum, washed, and then reacted with the hetero-antibody.

By use of specific antisera containing antibodies to a single immunoprotein, it is possible to discriminate which of the several immunoproteins is fixed to the membrane. Using such antisera, one can distinguish four types of immunoprotein coating on the red cells of patients with autoimmune hemolytic anemia:

Type I: immunoglobulin alone is present on the membrane.
Type II: both immunoglobulin and complement are present on the red cell membrane.
Type III: only complement components are present on the membrane.
Type IV: neither complement nor immunoglobulins is detectable on the red cell membrane.

In this review, I shall analyze each of the types of reactions as they relate to the pathogenetic mechanisms and to disease processes.

TYPE I—IMMUNOGLOBULIN ALONE IS FIXED TO THE RED CELL MEMBRANE

In order for this type of direct antiglobulin test to occur, two conditions must be met: (*a*) The immunoglobulin must be firmly fixed to the antigen at body temperature. If it is not, it will be washed away during the preparation of the cells for the antiglobulin test. (*b*) The first component of complement (C1) is not fixed by the antibody. This may occur either because the antibody is intrinsically unable to fix C1, or the antigen sites are too distant, so that two such molecules or molecular subunits cannot form the "doublet" necessary for the fixation of C1 (9). The former occurs in the case of IgG_4, IgA, and perhaps some IgM molecules (10), but the vast majority of autoimmune antibodies are IgM, IgG_1, and IgG_3, which are intrinsically capable of fixing C1 (11). The latter condition, placement of antigen sites, explains the majority of the instances in which complement is not fixed. Most autoimmune antibodies which do not fix complement have some specificity for the Rh system (12), and Rh antigens appear to be so dispersed that doublets cannot form (13).

Red cells coated with antibody which does not fix complement do not hemolyze in vitro but may be destroyed in vivo. This is thought to be due to activation of a part of the heavy chain portion of the antibody molecule to enable it to interact with specific sites on phagocytic cells (14, 15). This activation occurs only with IgG_1, IgG_2, and IgG_3, molecules upon interaction with antigen and does not occur in IgG_4, IgM, or IgA molecules (14, 16). Because of this interaction, the red cell is affixed to phagocytic cells, in particular to fixed macrophages in the spleen. It is either completely phagocytosed or a portion of its membrane is removed (17). In the latter case, the cell membrane is able to reseal and a spherocyte results; this abnormally shaped cell may then be trapped in subsequent circulation through the spleen (18). Thus, hemolysis of this type is characterized by relatively little evidence of intravascular lysis (increased plasma hemoglobin, hemoglobinuria) and often by the presence of spherocytes and splenomegaly.

With recent techniques which allow estimation of antibody concentration on red cells, it is found that the degree of hemolysis depends in part upon the concentration of antibody, but there is considerable variation in the amount of IgG attached for a given amount of hemolysis (19, 20).

In about one third to one fourth of all patients with AIHA in which IgG antibody is detectable on the cell membrane, no complement is fixed (2). This occurs in patients with chronic lymphocytic leukemia (CLL) as well as in patients with "idiopathic" immune disease. The antibodies of patients with systemic lupus erythematosus (SLE) and other "collagen vascular" diseases usually fix complement (12). Of the syndromes of drug-induced hemolysis, both that initiated by alphamethyldopa and penicillin are usually characterized by the fixation of antibody without the fixation of complement (21).

TYPE II—BOTH ANTIBODY AND COMPLEMENT ARE DETECTED ON THE RED CELL MEMBRANE

When the conditions for continued antibody attachment as well as those for fixation of C1 are met, both IgG and complement components (mostly C4 and C3) may be found on the red cell surface by the use of specific antisera (22). Thus, warm-reacting IgG_1, IgG_2, or IgG_3 antibodies of reasonably high affinity which react with plentiful antigens will give this reaction. These antibodies are encountered in "idiopathic" AIHA, and AIHA secondary to CLL and SLE, as well as other syndromes of AIHA. However, with some exceptions (23), the antibodies involved in drug-induced immune hemolysis do not fix both antibody and complement.

Hemolysis in vivo may occur both because of the activation of antibody (see above) or the activation of complement. Two mechanisms appear to be important in effecting hemolysis. 1. When the third component is affixed in its active form, a site on the molecule is able to fix to receptors on phagocytic cells, thus allowing phagocytosis (24, 25); this mechanism is "defused" by an enzyme which cleaves the C3 molecule, thus detaching its red cell binding site from its phagocyte binding site (26–28). 2. The fixation of the last component, C9, appears to breach the lipid bilayer, causing rupture of the cell (29, 30). When this occurs in vivo, intravascular hemolysis results. For unknown reasons, the normal red cell is difficult to breach by this mechanism; hence, large amounts of defused C3 and other components of complement may accumulate on the red cell surface (28).

If both antibody and complement were fixed, one would expect greater hemolysis for a given amount of antibody than if complement were not fixed. Careful studies in animals seem to confirm this (31) and studies are currently under way to determine, if possible, the relative roles of antibody and complement in effecting lysis in AIHA.

TYPE III—ONLY COMPLEMENT IS FIXED TO THE RED CELL SURFACE

In many instances of AIHA only the components of complement are found fixed to the red cell surface by the direct antiglobulin test using specific antisera. This may mean that the antibody has been there but was removed either in vivo or during the washing of the cells on preparation for the test. There are several reasons for this.

1. The antibody is "cold-reacting" and is thus eluted at the temperatures at which the test is performed, such as IgM cold agglutinins and the IgG Donath-Landsteiner antibody.

2. The antibody, while capable of reacting at 37° C, has a low affinity and is easily eluted during the washing of the cells. Since treatment of the red cells with proteolytic enzymes usually increases the affinity of antibody and antigen (32), such antibodies can often be detected in the serum using enzyme-treated cells in the indirect antiglobulin test.

3. The antigen is not firmly fixed to the red cell surface. In this instance, both antigen and antibody may be removed during the washing. If complement has been fixed, it is not removed and is detected. The best examples of unfixed antigens occur in certain instances of drug-induced immune hemolysis (the "innocent bystander" reactions) (21). This reaction may occur with antigens other than drugs, but this has not been proven.

In some instances of Type III reactions, antibody may be present but not detected by the antiglobulin serum for several reasons:

1. The antibody is not IgG. Most autoimmune warm-reacting antibodies are IgG but IgM and IgA antibodies are found, usually in conjunction with IgG antibodies (5, 6). These will not be detected by anti-IgG but will with specific antisera.

2. The antibody is present in too low concentration to be detected by the antiglobulin serum. The limits of the agglutination test for detection of antibody are discussed below.

3. The antibody is detected only at reduced temperature. Certain IgG antibodies are fixed to the red cell at 37° C but are not able to react with antiglobulin molecules unless the temperature is reduced (33). These antibodies frequently fix complement which is detectable on the red cell surface at any temperature.

The hemolysis which ensues in patients exhibiting Type III direct antiglobulin test is probably predominantly due to the presence of complement components on the surface. The amount of hemolysis may vary from relatively little to considerable.

TYPE IV—NEITHER ANTIBODY NOR COMPLEMENT IS DETECTED ON THE RED CELL MEMBRANE

In a few patients with what clinically appears to be AIHA, the direct antiglobulin test appears to be negative with both anti-IgG and anti-C3 (2). This may be due to the presence of amounts of antibody too small to be detected in the agglutination reaction with antiglobulin sera. Probably somewhere between 200 and 1000 molecules of antibody per red cell are required before the agglutination reaction is positive as usually performed (34), but amounts less than this can be detected by special techniques (35, 36). It should be remembered that the amount of antibody on the cell surface may be very much larger in vivo before the cells are washed in preparation for the antiglobulin test. Some patients with AIHA and a completely negative Coombs' test may have considerable hemolysis.

SUMMARY

The foregoing summarizes what can be learned from a carefully performed antiglobulin test using specific antisera. Clinical syndromes can be more easily considered, and mechanisms of destruction can be more certainly discerned.

Literature Cited

1. Pirofsky, B. 1969. *Autoimmunization and the Autoimmune Hemolytic Anemias,* Baltimore: Williams & Wilkins
2. Dacie, J. V., Worlledge, S. M. 1969. Auto-immune hemolytic anemias. *Progr. Hematol.* 6:82–120
3. Ishizaka, T., Ishizaka, K., Borsos, T., Rapp, H. 1966. C'1 fixation by human isoagglutinins: fixation of C'1 by λG and λM but not by λA antibody. *J. Immunol.* 97:716–26
4. Yount, W. J., Dorner, M. M., Kunkel, H. G., Kabat, E. A. 1968. Studies on human antibodies. VI. Selective variations in subgroup composition and genetic markers. *J. Exp. Med.* 127:633–46.
5. Hsu, T. C. S., Rosenfield, R. E., Burkart, D., Wong, K. Y., Kochwa, S. 1974. *Vox Sang.* 26:305
6. Engelfriet, C. P., Von dem Borne, A. E., Van der Geissen, M., Beckers, Do, Van Loghem, J. J. 1968. Autoimmune haemolytic anaemias. I. Serological studies with pure anti-immunoglobulin reagents. *Clin. Exp. Immunol.* 3:605–14
7. Dacie, J. V. 1962. *The Haemolytic Anaemias, Part II, The Auto-Immune Haemolytic Anaemias.* New York: Grune & Stratton
8. Coombs, R. R. A., Mourant, A. E., Race, R. R. 1945. A new test for the detection of weak and "incomplete" Rh agglutinins. *Brit. J. Exp. Pathol.* 26:255–66
9. Rosse, W. F. 1968. Fixation of the first component of complement C'1a) by human antibodies. *J. Clin. Invest.* 47:2430–45
10. Ishizaka, T., Ishizaka, K., Borsos, T., Rapp, H. 1966. C'1 fixation by human isoagglutinins: fixation of C'1 by λG and λM but not by λA antibody. *J. Immunol.* 97:716–26
11. Borsos, T., Rapp, H. J. 1965. Complement fixation on cell surfaces by 19S and 7S antibodies. *Science* 150:505–6
12. Leddy, J. P., Peterson, P., Yeaw, M. A., Bakemeier, R. F. 1970. Patterns of serologic specificity of human λG erythrocyte autoantibodies. Correlation of antibody specificity with complement-fixing behavior. *J. Immunol.* 105:677–8
13. Nicholson, G. L., Masouredis, S. P., Singer, S. J. 1971. Quantitative two-dimensional ultrastructural distribution of Rh_o (D) antigenic sites on human erythrocyte membranes. *Proc. Nat. Acad. Sci. USA* 68:1416–20
14. Abramson, N., Gelfand, Jandl, J. H., Rosen, F. S. 1970. The interaction between human monocytes and red cells. Specificity for IgG subclasses and IgG fragments. *J. Exp. Med.* 132:1207–15
15. Abramson, N. Lo Buglio, A. F., Jandl, J. H., Cotran, R. S. 1970. The interaction between human monocytes and red cells. Binding characteristics. *J. Exp. Med.* 32:1191–1206
16. Abramson, N., Schur, P. H. 1972. The IgG subclasses of red cell antibodies and relationship to monocyte binding. *Blood* 40:500–8
17. Lo Buglio, A. F., Cotran, R. S., Jandl, J. H. 1967. Red cells coated with immunoglobulin G: Binding and sphering by mononuclear cells in man. *Science* 158:1582–85
18. Weed, R. I., Reed, C. 1966. Membrane alteration leading to red cell destruction. *Am. J. Med.* 41:681–98
19. Rosse, W. F. 1971. Quantitative immunology of immune hemolytic anemia. II. The relationship of cell-bound antibody to hemolysis and the effect of treatment. *J. Clin. Invest.* 50:734–43
20. Constantoulakis, M., Costea, N., Schwartz, R. S., Dameshek, W. 1963. Quantitative studies of the effect of red-blood-cell sensitization on in vivo hemolysis. *J. Clin Invest.* 42:1790–1801
21. Worlledge, S. M. 1973. Immune drug-induced hemolytic anemias. *Semin. Hematol.* 10:327
22. Harboe, M., Muller-Eberhard, H. J., Fudenberg, H., Polley, M. J., Mollison, P. L. 1963. Identification of the components of complement participating in the antiglobulin reaction. *Immunology* 6:412–20
23. Kerr, R. O., Cardamone, J., Dalmasso, A. P., Kaplan, M. E. 1972. Two mechanisms of erythrocyte destruction in penicillin-induced hemolytic anemia. *N. Engl. J. Med.* 287:1322–25
24. Lay, W. H., Nussenzweig, V. 1968. Receptors for complement on leukocytes. *J. Exp. Med.* 128:991–1009
25. Gigli, I., Nelson, R. A. Jr. 1968. Complement dependent immune phagocytosis. I Requirements for C'1, C'4, C'2, C'3. *Exp. Cell Res.* 51:45–67
26. Ruddy, S., Austen, K. F. 1971. C3b inactivator of man. II. Fragments produced by C3b inactivator cleavage of

cell-bound or fluid phase C3b. *J. Immunol.* 107:742–50

27. Ruddy, S., Austen, K. F. 1969. C3 inactivator of man. I. Hemolytic measurement by the inactivation of cell-bound C3. *J. Immunol.* 102:533–43
28. Logue, G. L., Rosse, W. F., Adams, J. P. 1973. Complement-dependent immune adherence measured with human granulocytes: Changes in the antigenic nature of red cell bound C3 produced by incubation in human serum. *Clin. Immunol. Immunopathol.* 1:398–407
29. Mayer, M. M. 1972. Mechanism of cytolysis by complement. *Proc. Nat. Acad. Sci. USA* 69:2954–58
30. Kolb, W. P., Haxby, J. A., Arroyave, C. M., Müller-Eberhard, H. J. 1972. Molecular analysis of the membrane attack mechanism of complement. *J. Exp. Med.* 135:549–66
31. Schreiber, A. D., Frank, M. M. 1972. Role of antibody and complement in immune clearance and destruction of erythrocytes. *J. Clin. Invest.* 51:575–82
32. Hughes-Jones, N. C., Gardner, B., Telford, R. 1964. The effect of ficin on the reaction between anti-D and red cells. *Vox Sang.* 9:175–82
33. MacKenzie, M. R., Creevy, N. C. 1970. Hemolytic anemia with cold detectable IgG antibodies. *Blood* 36:549–58
34. Dupuy, M. E., Elliot, M., Masouredis, S. P. 1964. Relationship between red cell bound antibody and agglutination in the antiglobulin reaction. *Vox Sang.* 9:40–44
35. Gilliland, B. C., Leddy, J. P., Vaughn, J. H. The detection of cell-bound antibody on complement-coated human cells. *J. Clin. Invest.* 49:898–906
36. Burkart, P., Rosenfield, R. E., Hsu, T. C. S., Wong, K. Y., Kochwa, S. 1974. *Vox Sang.* 26:289

Copyright 1975. All rights reserved

IMMUNOTHERAPY (DESENSITIZATION) IN ALLERGIC DISEASE[1]

❖7155

Philip S. Norman, M.D.
Department of Medicine, The Johns Hopkins University; and Allergy Service, The Good Samaritan Hospital, Baltimore, Maryland 21239

The practice of therapeutic immunization for allergic conditions (desensitization, immunotherapy) was introduced in 1911 (1) and within a decade became widely accepted on an empirical basis. As usually practiced by specialists in allergy, the technique consists of selecting those naturally occurring allergens such as pollens, molds, animal danders, dusts, etc, which by clinical history and diagnostic skin testing appear to be responsible for the immediate type of allergic symptoms. A mixture of crude extracts specific for the individual patient is prepared and a series of injections is initiated with a high dilution, gradually increasing the strength of the dilution at weekly or biweekly intervals. Dosage is gradually increased in this fashion until a predetermined maximum dosage is administered with each injection, or the highest tolerated dose, as estimated by local or systemic allergic reactions to the injections, is reached. At this point maintenance injections are continued at intervals from two to four weeks. Results are commonly assessed by questioning the patient about symptoms at the time of exposure to allergens.

After many years of such empirical use, in the last ten years an organized collection of evidence has produced a sounder basis for judging the efficacy of the process in allergic conditions. At the same time immunologic study provided some important, although incomplete, insights into the mechanism of the partial relief of symptoms provided by immunotherapy. This information provides a rational basis for the use of the method in certain instances, although at this time the data are still sufficiently incomplete that the clinician must continue to rely on individual judgment or traditions built up by many years of practical use as the basis for decision in many, and perhaps most, cases.

[1] Publication No. 153 of the O'Neill Memorial Research Laboratories.

In this review, immunologic information will be outlined in order to develop current knowledge of the rationale for immunotherapy, following which those systematic and controlled observations that bear on the clinical efficacy of the method will be discussed. The resulting judgments are at best tentative and the pace of continuing investigations is such that many changes in these judgments may be called for in the next few years.

IMMUNOLOGIC RESPONSES

Current knowledge of the pathogenesis of immediate allergic reactions depends on the discovery that "skin sensitizing antibodies" or "reagins," which are demonstrated by passive transfer in the skin of normal volunteers by the Prausnitz-Küstner test, belong to a special class of antibodies known as immunoglobulin E (IgE) (2). IgE antibodies are synthesized principally in lymphoid tissue associated with the respiratory and gastrointestinal tracts and circulate in the blood in extremely minute amounts (0.1–0.7 μg/ml serum) but have a strong affinity for surface receptors on tissue mast cells and basophils. Once IgE antibodies to an allergen have attached to the membranes of such cells, the cells are "sensitized" and will respond to contact with the allergen with a secretion of histamine, slow reacting substance-A (SRS-A), and probably other biochemical mediators that produce the manifestations of allergic reactions. In allergic individuals the synthesis of IgE antibodies continues for unusually long periods by ordinary immunologic standards. In the case of ragweed hay fever, an annual exposure of the respiratory mucosa to extremely small amounts of ragweed pollen allergen (estimated to be only a few micrograms) is sufficient to restimulate IgE antibody production to that antigen so that blood levels are maintained throughout the following year with only a slight decrease by the time of the next seasonal exposure (3).

The severity of disease upon respiratory exposure to an allergen such as ragweed appears to depend upon the amount of IgE antibodies to ragweed attached to the membranes of mediator-containing cells which, in turn, is related to the amount of circulating anti-ragweed IgE antibody (4). Mediator-containing cells can be sampled directly by making a suspension of viable white cells from freshly drawn samples of blood. Reactions of mediator release from the basophils in the sample can be tested in vitro by exposing the cell suspension to allergen at physiological conditions of pH, temperature, ionic strength, glucose concentration, and divalent cation concentration. Reactions are measured by chemical determination of the histamine released (5).

IgE antibodies can be measured in serum specimens by the so-called radioallergosorbent test (RAST), which depends on coupling the antigen to an insoluble particle, and then measuring the amount of IgE antibody that combines with the allergen by a reagent which is an antibody to human IgE raised in another species and then radiolabelled (4).

Although immunologic responses to immunotherapy were studied by the early workers by the relatively crude biologic passive transfer techniques then available and some insights were obtained, the more recently developed in vitro techniques described above have allowed a more precise definition of immunologic responses to immunotherapy. Early work demonstrated the development after immunization of "blocking antibodies," a different type of heat stable antibody than the reagins detected by passive transfer. The development of the in vitro histamine release technique allowed the demonstration that every allergic individual, even those not yet subjected to perennial immunization, carries small amounts of blocking antibodies in the serum, such antibodies being demonstrated by their ability to inhibit basophil histamine release (6). Studies of individuals undergoing immunization for ragweed hay fever showed that levels of such blocking antibodies increased as the dosage of the injections was gradually increased (7). These antibodies do not fix to mediator-containing cells but rather circulate freely, and are of the IgG class. When injections are discontinued, levels of such IgG blocking antibodies decline in a matter of months but can be restimulated by reinstitution of the injections.

IgE Antibodies

By several different techniques it has been demonstrated that early in the course of an injection series, when the doses are very low, there is actually a slight increase in the specific IgE antibody. However, as injections continue, there may actually be a modest decrease in the level of specific IgE antibody for the allergen being injected. What is of particular interest, however, is that at the time of annual seasonal exposure by the respiratory route the usual increase of IgE antibodies toward the allergen is considerably blunted. Thus, after a number of years of repeated injections, the level of IgE antibody may be noticeably reduced (3). In a number of patients, however, when injections are discontinued, the next seasonal exposure to allergen results in a steep rise in specific IgE antibodies (2).

Changes in Cellular Responses

Another change, which may not be immunologic, has been observed in basophil reactivity. Some patients require increasing amounts of antigen to cause basophil histamine release as the injections progress. This change may progress to the point where only partial release of the cellular store of histamine occurs even at the highest concentration of antigen possible, and in the extreme instance virtually no histamine release occurs. The reduction in reactivity appears to be nonspecific, as the basophils are incapable of releasing histamine even to allergens not included in the patient's treatment mixture. Basophil reactivity tends to return when injections are discontinued (8).

CLINICAL EFFICACY

At the same time that the study of immunologic change has been undertaken, a fairly extensive controlled study of the clinical efficacy of desensitization regimens has been carried out. In some instances clinical results and immunologic changes have been examined in a correlative way in the same patients. The greater proportion of these studies has been carried out in patients with seasonal hay fever, as this condition is relatively common and represents a well-circumscribed disease amenable to study.

Pollen Disease

Controlled double-blind studies in ragweed and grass hay fever comparing a preseasonal desensitization regimen with placebos demonstrate that the treated individuals undergo a considerable reduction in the severity of symptoms in many but not all instances. Evaluation of symptom relief depends either on structured interviews by physicians *during* the course of the season of exposure or on meticulously prepared diaries of symptoms on structured reporting forms. Either method can be scored numerically, and the scores of groups of specifically treated and placebo-treated patients compared and differences tested by appropriate statistical methods. Such scoring methods show that relatively few cures occur and most treated individuals suffer at least some degree of symptoms (9, 10).

In ragweed hay fever a significant but rather rough correlation between the titer of blocking antibody to ragweed which develops and symptom relief has been demonstrated. This correlation is found only in groups of patients studied and the levels of blocking antibody have little predictive value in individual patients. In a general way, however, patients with the highest level of blocking antibody have the greatest symptom relief (11). It was clear, however, again on the basis of groups of patients, that the degree of relief depended on dose and the groups who received the higher doses of extract fared better clinically. The upper level of dosage achievable was, however, usually limited by the appearance of allergic reactions to the injections given, so that attempts to raise the dose still further would have resulted in an inordinate number of injections and an increasing danger of unpleasant, or even life-threatening, allergic reactions (12). In the case of ragweed extract, comparison of results using a purified antigenic principal protein known as "Antigen E" with results with whole extracts strongly suggested that the major portion of the therapeutic efficacy of ragweed extracts depended on the presence of Antigen E in the extract (10).

Although some of the changes in cellular reactivity appeared to be nonspecific, the clinical results of immunotherapy where studied appear to be highly specific. Patients who have both ragweed and grass hay fever, when treated with ragweed extract, only show relief of ragweed-induced symptoms but suffer unabated grass hay fever. In addition, patients treated successfully for four seasons showed a high rate of relapse in the two seasons following the discontinuance of injections (13). This is consistent with the fall of IgG-blocking antibodies after discontinuance and the return of the IgE antibodies at the time of next exposure.

These studies were performed in adults. Additional study in groups of children tended to confirm that the same principles applied to treatment of children, in that the quality of results depended on dosage and cures were infrequent. The degree of symptom relief in children seemed to be better than in adults when looked at by these methods of evaluation, thus confirming rather long standing clinical impressions (8).

Relief from seasonal asthma apparently due to pollen exposure has been similarly studied clinically, although without benefit of simultaneous immunologic investigation. Good symptom relief was obtained in both children and adults, and some patients seemed cured (9, 14).

House Dust Allergy

Types of allergy other than those due to seasonal pollen exposure are less completely studied. Allergy to house dust causing asthma has been studied with mixed results. One cooperative study in England using rather small doses over a comparatively short time of only 15 injections revealed no improvement of the patient's asthma (15). However, a study of bronchial sensitivity after more extensive injections of house dust extract showed a reduction in bronchial responses to inhaled extract of house dust (16). A long term study over several years in children showed that two different house dust extracts given over a period of two and a half to three years produced a greater rate of improvement clinically, accompanied by a reduction in bronchial sensitivity, than was demonstrated in the similar group of children given placebos (17). The discovery that mites living in house dust represent a major allergen, at least in northern humid climates, has led to studies of desensitization in house dust sensitive patients with extracts of mites in areas where house dust mites seem to be of great importance. Such studies have indicated a significant amelioration of asthmatic symptoms (18). Obviously, one of the problems with house dust extract is the heterogeneous nature of house dust and the likelihood that different patients are sensitive to different components. Histamine release studies on basophils from house dust sensitive patients show that several commercially prepared dust extracts contain amounts of allergenic material that were quite small in comparison with the usual kind of pollen extract. Thus, the protein dosage, in order to achieve blocking antibody production with dust extract, would have to be much larger than it ordinarily is currently with house dust or pollen extracts (19). There is no clinical or immunological study of the effects of extracts of molds or animal danders to date so that it is still not known whether the principles learned with pollen extracts will apply to these substances.

Bacterial Vaccines

An entirely different area is the use of killed bacterial vaccines in desensitization regimens in patients whose rhinitis or asthmatic symptoms seem to arise at the time they have respiratory infections. The rationale here is that the patients have become sensitive to the bacteria that participate in these infections. The vaccines are commonly stock mixtures containing organisms found in the mouth and throat such as pneumococcus, streptococcus, staphylococcus, *Neisseria, Klebsiella,* and *Hemophilus.* There is virtually no evidence that these are the organisms that produce the respiratory infections these patients develop. Indeed, all evidence currently available would suggest that most self-limited respiratory infections in both children and adults arise from respiratory viruses. Nevertheless, the use of such vaccines either as the sole agent in the desensitization regimen or as an addition to a regimen which includes pollens, molds, and danders has become a common practice on the basis that at least some of the patients so treated improve insofar as the frequency and severity of asthmatic attacks associated with respiratory infections are concerned.

A number of controlled studies have been carried out to test whether use of one or another variation of these vaccines actually helps patients more than placebos. There are six such studies and while each can be criticized on the basis of selection of patients, adequacy of blind techniques, sample size, and dependence on impressions by physicians, there has been an almost complete lack of positive results. One must conclude that, while prevention or early cure of respiratory infections in patients with asthmatic disease during such infections should be a useful maneuver, such bacterial vaccines have little ability to prevent the recurrence of such infections or the development of an asthmatic reaction when such infections occur (13).

Clinical Conclusions

A properly applied regimen of immunotherapy is sufficiently effective in pollen disease and probably in mold- and animal dander-induced conditions to deserve a place in the physician's armamentarium for the treatment of these conditions. Nevertheless, in the therapeutic decision in any individual case the partial effectiveness of immunotherapy and the hazards of anaphylactic reaction must be balanced against the benefit and side effects that may occur from the use of drug therapy. In hay fever the physician has available to him antihistamines, decongestants, and topically applied corticosteroids in the nose. The latter, in particular, gives a high rate of relief with little or no steroid side effects. For asthma, bronchodilator drugs and systemically administered corticosteroids will provide relief, although the dangers of corticosteroids are well known. For simple seasonal hay fever, which will continue only for a limited period, drug therapy may often be more satisfactory from the standpoint of convenience and low cost. For patients with multiple or perennial allergies or asthmatic disease, it seems appropriate from the beginning to employ an immunization program along with symptomatic treatment by drugs where needed.

NEWER METHODS

Because the results of immunotherapy are dose-related and seem to be increasingly successful as the dose increases, and yet the dose is limited because of untoward allergic reactions, various techniques have been sought to employ larger doses safely. The first approach has been to slow absorption by an additive that can restrict solubility. Alum precipitation and emulsification in oil have been the two methods most often tried. Oil emulsions have been tried most extensively but have not always been successful clinically (20). A serious question as to the safety of the mineral oil usually employed has been raised by the ability of large doses of oil to induce myelomas in mice (21). Hence approval of a product containing mineral oil for widespread use seems unlikely. Alum precipitation makes a less satisfactory repository but has no known problem in regard to safety. Alum-precipitated ragweed extract can be given with modest increases in dosage in about one half the usual number of injections with a somewhat lowered rate of systemic allergic reactions

(22). Such extracts are available to the physician for a variety of allergens. They should be distinguished from extracts made with the organic solvent, pyridine, which are then alum-precipitated. Pyridine has been shown to denature some of the essential allergens in at least ragweed extract (23). Efforts to couple allergenic materials to various insoluble or poorly soluble carriers are under way in several laboratories to make a better repository (24).

Another approach to modifying extracts for therapy is chemical modification to change their ability to induce antibody production. The term "allergoid' has been used to describe formaldehyde-modified allergens. By treatment with formaldehyde under mild conditions, it has been possible to prepare grass and ragweed pollen allergoids which have greatly reduced ability to trigger allergic response, as measured by either skin reactions or basophil histamine release, but which can nevertheless elicit IgG antibody responses in good titer (25). Preliminary trials in allergic patients indicate that administration of allergoids is safe and does elicit blocking antibody in good titer. Clinical efficacy is now being tested.

STANDARDIZATION

A continuing problem with the aqueous extracts most commonly employed by allergists has been adequate standardization. Labelling of such extracts is usually based either on simple weight of allergen extracted over volume of extracting fluid or on a total protein determination. It is readily apparent from identification and purification of allergenic moieties of ragweed, grass, cottonseed, and fish that the proteins responsible for symptoms and also for the desirable immunologic and therapeutic effects of injections of extract represent only a small proportion of the total protein in the pollen. Indeed, in the case of ragweed extract, the principal allergen, Antigen E, represents only 0.5% of the extractable solids and about 6% of the extractable protein (26). The total protein determination, therefore, does not necessarily reflect the actual content of the proteins of interest and importance in diagnosis and therapy. The problem is compounded by the lability of some of these under even mildly adverse conditions. Indeed, when ragweed extract with the same label in terms of protein was purchased from six different manufacturers, the biologic potency as measured by skin tests and histamine release, and the Antigen E content as measured by immunologic methods, varied by 100-fold (27). Similar evidence has been obtained with grass pollen extracts. Once the principal allergenic protein has been identified and purified, it is possible by immunologic methods to determine quantitatively its presence in any extract and to study methods of stabilization and determine shelf life of the active principal. However, the allergenic proteins have been identified and purified in only a limited number of allergens so that that method is not yet universally applicable. Again, work in several laboratories is directed toward using some modification of the specific in vitro IgE reaction to determine the potency of allergenic extracts in a quantitative fashion.

It will be apparent from this review that the combination of controlled clinical study and new immunologic methods has served to confirm some clinical impressions based on long practice and to bring other empirical findings into question. Out of such investigation it can be hoped that improved methods for the prophylaxis of allergic symptoms will be developed. The rapid advance of immunologic knowledge seems likely to disclose still other avenues for intervention in the abnormal processes which lead to allergic symptoms. Although one can only conjecture as to the nature of such intervention, the developing picture of cooperation between B and T lymphocytes in immunologic responses may offer additional opportunities for the specific direction of antibody responses in a fashion that would safely eliminate allergic responses without sacrifice of the protective function of the immunologic body defenses against infections.

Literature Cited

1. Noon, L. 1911. *Lancet* 1:1572–73
2. Ishizaka, K., Ishizaka, T. 1973. *Asthma, Physiology, Immunopharmacology and Treatment,* ed. K. F. Austen, L. M. Lichtenstein, Chap. 5, 55–68. New York: Academic
3. Lichtenstein, L. M., Ishizaka, K., Norman, P. S., Sobotka, A. K., Hill, B. M. 1973. *J. Clin. Invest.* 52:472–82
4. Norman, P. S., Lichtenstein, L. M., Ishizaka, K. 1973. *J. Allergy Clin. Immunol.* 52:210–24
5. Lichtenstein, L. M., Osler, A. G. 1964. *J. Exp. Med.* 120:507–30
6. Lichtenstein, L. M., Osler, A. G. 1966. *J. Immunol.* 96:169–79
7. Lichtenstein, L. M., Norman, P. S., Winkenwerder, W. L., Osler, A. G. 1966. *J. Clin. Invest.* 45:1126–36
8. Levy, D. A., Lichtenstein, L. M., Goldstein, E. O., Ishizaka, K. 1971. *J. Clin. Invest.* 50:360–69
9. Frankland, A. W., Augustin, R. 1954. *Lancet* 1:1055–57
10. Norman, P. S., Winkenwerder, W. L., Lichtenstein, L. M. 1968. *J. Allergy* 42:93–108
11. Lichtenstein, L. M., Norman, P. S., Winkenwerder, W. L. 1971. *Ann. Intern. Med.* 75:663–71
12. Norman, P. S., Winkenwerder, W. L., Lichtenstein, L. M. 1971. *J. Allergy* 47:273–82
13. Norman, P. S. 1974. *Med. Clin. N. Am.* 58:111–25
14. Johnstone, D. E. 1957. *Am. J. Dis. Child.* 94:1–5
15. Forgacs, P., Swan, A. V. 1968. *Brit. Med. J.* 3:774–77
16. McAllen, M. K. 1961. *Thorax* 16:30–35
17. Aas, K. 1971. *Acta Pediat. Scand.* 60:264–68
18. Smith, A. P. 1971. *Brit. Med. J.* 4:204–6
19. May, C. D., Lyman, M., Alberto, R., Cheng, J. 1970. *J. Allergy* 46:73–89
20. Norman, P. S., Winkenwerder, W. L., D'Lugoff, B. C. 1967. *J. Allergy* 39:89–92
21. Potter, M., Boyce, C. R. 1962. *Nature* 193:1086–87
22. Norman, P. S., Winkenwerder, W. L., Lichtenstein, L. M. 1972. *J. Allergy Clin. Immun.* 50:31–44
23. Lichtenstein, L. M., Norman, P. S., Winkenwerder, W. L. 1968. *J. Allergy* 41:49–57
24. King, T. P., Norman, P. S., Tao, N. 1974. *Immunochemistry* 11:83–92
25. Marsh, D. G., Lichtenstein, L. M., Norman, P. S. 1972. *Immunology* 22:1013–28
26. King, T. P., Norman, P. S., Connell, J. T. 1964. *Biochemistry* 3:458–68
27. Baer, H., Godfrey, H., Maloney, C. J., Norman, P. S., Lichtenstein, L. M. 1970. *J. Allergy* 45:347–54

Copyright 1975. All rights reserved

THALASSEMIA

◆7156

Bernard G. Forget, M.D. and David G. Nathan, M.D.
The Division of Hematology and Oncology of the Department of Medicine, Children's Hospital Medical Center, the Sidney Farber Cancer Center, and the Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115

INTRODUCTION

The thalassemia syndromes constitute a heterogeneous group of hereditary disorders of human hemoglobin synthesis found largely in individuals of Mediterranean, African, and Oriental ancestry. These disorders are characterized by absent or decreased synthesis of the alpha (α) or beta (β) globin chain of human adult hemoglobin, Hb A [$\alpha_2\beta_2$] (1, 2). In those cases where some of the affected globin chain is synthesized, there is no evidence of an amino acid substitution. However, in all cases where genetic evidence is available, the thalassemia gene appears to be allelic to the structural gene for the α or β globin chain.

In this review we will attempt to summarize current concepts of the pathophysiology, molecular basis, clinical diagnosis, and management of the more common forms of thalassemia.

PATHOPHYSIOLOGY

An immediately obvious consequence of the decreased synthesis of one of the two subunits of Hb A is an overall deficit in complete Hb A tetramers accumulated within each erythrocyte. A hypochromic microcytic anemia results.

In the more severe forms of thalassemia, such as homozygous β-thalassemia, and the α-thalassemia syndrome, Hb H disease, there is, in addition to severe hypochromic microcytosis, a moderate to severe hemolytic process which results from the marked imbalance in globin chain synthesis occurring in the erythroid cells. In these conditions, the continued synthesis, in normal amounts, of the nonaffected globin chain, (α chains in β-thalassemia and β chains in hemoglobin H disease) leads to the accumulation of a large relative excess of these normal chains within the red cell. Not finding complementary globin chains with which to bind, these normally synthesized globin chains bind to one another and form aggregates that precipitate within the cell and become attached to the cell membrane. These precipitates damage the red cell membrane: the red cells become more permeable to cations, become depleted of ATP, and are prematurely destroyed in the reticuloendothelial system (3). The reticuloendothelial system, especially in the spleen and marrow itself, is very active and effective in removing the inclusion bodies (globin chain aggregates) that form and precipitate in the red cells. This process of "pitting" of inclusions from newly formed and circulating red cells is also responsible for ineffective erythropoiesis, shortening of red cell survival, and hemolysis. In the process of removal of inclusion bodies, the reticuloendothelial cells must pinch off a portion of the red cell membrane; the membrane reseals, and the cell may return, deformed, to the circulation but it is likely, after one or more such pitting episode, to be irreparably damaged and destroyed in the process.

In homozygous β-thalassemia, the imbalance between α and β chain synthesis can, in some cells, be partially corrected by increased synthesis of the γ chain of fetal hemoglobin, Hb F. Hb F is heterogeneously distributed among the red cells of these patients: those cells with abundant Hb F have fewer α chain inclusion bodies and are longer lived, whereas the cells with less Hb F have many more inclusion bodies and a shorter survival (1, 3).

MOLECULAR BASIS OF THALASSEMIA

The primary molecular mechanisms responsible for the imbalance of globin chain synthesis in thalassemia are not known. However, because of recent technological advances in the fields of biochemistry and molecular biology, which have been applied to research on thalassemia, there has emerged a more precise concept of the nature of the possible molecular defect(s) in thalassemia. It has been found that thalassemic erythroid cells lack normal amounts of globin messenger RNA (mRNA) for the affected globin chain. The precise cause of this deficiency is unknown but a limited number of possibilities exist. This subject has been recently reviewed in detail (2).

The role of globin mRNA in the thalassemic process was first directly demonstrated when it became possible first to isolate and then translate globin mRNA in various heterologous cell-free protein synthesizing systems. Cell-free translation of mRNA isolated from reticulocytes of patients with homozygous β-thalassemia resulted in the same imbalance of globin chain synthesis as occurs in the intact β-thalassemic reticulocytes: marked decrease of β chain synthesis relative to α chain synthesis (4, 5). Similar studies have since been repeated with mRNA isolated from reticulocytes and marrow cells of a number of patients with various thalassemia syndromes (2).

True quantitative deficiency of chain-specific globin mRNA has been subsequently demonstrated in thalassemic reticulocyte RNA by three different techniques which permit quantitation of the relative amounts of α and β chain specific mRNAs, independently of their capacity to function as a template for protein synthesis: 1. DNA-RNA hybridization assays which utilize, as probes, radioactive complementary DNA (cDNA) synthesized, from purified α and β specific mRNA, by means of the RNA-dependent DNA polymerase of avian myeloblastosis virus (6, 7); 2. direct nucleotide sequence analysis (fingerprinting) of thalassemic mRNA

(8); and 3. analysis of thalassemic mRNA by polyacrylamide gel electrophoresis in the presence of formamide which separates α and β mRNAs (27).

In all cases of thalassemia where definitive sensitive techniques of DNA-RNA hybridization or fingerprint analysis have been applied, quantitative deficiency of chain-specific globin mRNA has been demonstrated (2). Quantitative deficiency of α or β globin mRNA is therefore a common defect in most if not all the various thalassemia syndromes. However, different primary causes for this mRNA deficiency are likely to exist in different specific syndromes: 1. in those cases where there is absence of α or β chain synthesis and absence of α or β mRNA, a gene deletion is a likely cause; in fact such a deletion has been demonstrated in the homozygous α-thalassemia syndrome of hydrops fetalis with Hb Bart's (9, 10); 2. in those cases where some α or β chains are being synthesized, but in reduced amounts, and when there is some, but decreased amounts of the affected mRNA, at least two possibilities exist: (*a*) decreased transcription (synthesis) of the mRNA from the globin gene DNA, or (*b*) initial synthesis of the mRNA in normal amounts, but accelerated degradation in nucleated precursors, due to abnormal processing or transport. There is as yet no experimental evidence in favor of the presence or absence of any of these latter processes in thalassemia.

DIAGNOSIS

Heterozygous α- or β-thalassemia (thalassemia trait or thalassemia minor) are the most commonly encountered thalassemic conditions. They are characterized by microcytosis and hypochromia of the red blood cells, and a blood smear which reveals mild to moderate aniso- and poikilocytosis. There is usually only a mild anemia (Hb of 11–12 g%). The adequate Hb level is achieved at the expense of an abnormally high RBC count. With the recent availability of electronic blood cell counters, the diagnosis of thalassemia trait may be frequently made by the incidental finding of microcytosis (low mean corpuscular volume (MCV)) on a routine blood count (11). The condition most frequently confused with thalassemia trait is iron deficiency anemia, which is also characterized by hypochromia and microcytosis. In the absence of the definitive tests (serum Fe, iron-binding capacity (IBC), and Hb electrophoresis), the two conditions may be difficult to distinguish one from another, although the degree of microcytosis (12–14) and cell deformity is more extreme in thalassemia than in iron deficiency at Hb levels of 10–12 g%.

Ultimately Hb electrophoresis and serum Fe/IBC levels (or bone marrow aspiration for iron stain) offer the correct diagnosis. In heterozygous β-thalassemia the Hb A_2 will be elevated and constitute between 3.5 to 5% of the total Hb (normal Hb A_2, 2–3%) with or without a slight elevation of Hb F; in some rare cases (so-called F-thalassemia or $\delta\beta$-thalassemia), Hb A_2 may be normal but Hb F is elevated in the range of 5–15%. It should also be noted that in cases of concomitant iron deficiency with β-thalassemia trait, the Hb A_2 may fall to normal levels, but it will rise again to elevated levels after the iron deficiency is treated (15). In heterozygous α-thalassemia, the Hb A_2 and F levels are normal. The diagnosis is often one of exclusion after iron deficiency and other hypochromic anemias have been ruled out by appropriate tests. The importance of establishing the diagnosis of thalassemia trait is twofold: 1. to provide proper genetic counseling and, probably more important, 2. to avoid inappropriate treatment of the patient with medicinal iron in a long-term, and potentially harmful, fashion.

MANAGEMENT

The management of homozygous β-thalassemia is essentially supportive, and aimed at correcting the anemia and treating the complications associated with the progression of the disease, mainly the consequences of chronic iron overloading.

Transfusion Therapy

In the past, transfusions were given as infrequently as possible in order to delay as long as possible the sequelae of transfusion hemosiderosis. As a result, patients with homozygous β-thalassemia were chronically incapacitated. More recently, it has become obvious, through clinical studies started by Wolman et al (16) and undertaken in a number of other centers (17, 18), that patients do much better if their Hb is maintained at more physiological levels. In the so-called hyper- or high transfusion programs, patients are transfused as frequently as is necessary to maintain a minimum Hb level of 9–10 g%. As a result, the patients lead more normal lives and activities, have near normal exercise tolerance, develop fewer intercurrent illnesses. Certain complications, usually associated with thalassemia, are also less likely to develop or are less severe: hepatosplenomegaly and hypersplenism, mongoloid facies and associated orthodontic problems, marked bone marrow expansion, and recurrent fractures. The transfusions, by maintaining near normal Hb levels, suppress the erythropoietic stimulation of the marrow sufficiently to prevent overexpansion of the marrow and limit extramedullary hematopoiesis. There is not yet agreement on whether the usual growth retardation seen in thalassemia may be prevented by this type of transfusion program. Even though such transfusion programs result in the infusion of greater quantities of red cell iron than a more conventional program, the longevity of the patients, although not improved, does not appear to be shortened due to earlier or more severe hemosiderosis (16).

Chelating Agents

Ideally one would like to be able to remove some or most of the iron which accumulates due to the repeated transfusions, especially in patients being treated with high transfusion programs. The most commonly used iron chelating agent is desferrioxamine (Desferal®), at a dose of 0.5–1.0 g i.m. daily. This agent, unfortunately, requires daily intramuscular injections and past studies have shown that its effectiveness usually decreases substantially in time so that relatively little iron is excreted on a chronic basis compared to the input. More recently, the combination of 0.2–1.0 g of ascorbic acid orally per day with Desferal has been shown to result in a sustained increase in the urinary excretion of iron to levels where near balance may be achieved between iron infusion and urinary iron excretion (19, 20).

It should be pointed out, however, that Desferal will not prevent the development of transfusion hemosiderosis, even if given regularly from an early age (20). In fact, Desferal is not effective unless the total body iron stores are already increased to at least three times normal, i.e. only after hemosiderosis is already established (19, 20). It seems to be most effective in older patients with the highest body burden of iron. In some patients, the chronic use of Desferal is thought to have prevented the progression of hepatic fibrosis, which is presumably due to iron overload, and actually caused the liver iron content to decrease (21). Therefore, in selected patients shown to be responsive to it, Desferal may be a useful adjunct to a high transfusion program.

A new, orally effective chelating agent, dihydroxybenzoic acid, is currently being tested (22) and hopefully may prove to be more generally effective and acceptable.

Splenectomy

Splenectomy has long been a useful therapeutic modality, especially in patients with marked splenomegaly and a high transfusion requirement. In such patients, the spleen aggravates the anemia in two ways: 1. by creating a dilutional anemia because of its sheer mass which acts as a third space; and 2. by becoming truly hypersplenic and selectively sequestering and destroying red blood cells. A useful procedure to detect these phenomena and predict a good response to splenectomy is a ^{51}Cr red blood cell survival study with scanning of the spleen for detection of ^{51}Cr sequestration. Patients with short survival of homologous blood (^{51}Cr half time of 10–15 days) and with spleen/liver ^{51}Cr ratios of greater than 3:1 will usually benefit from splenectomy and achieve a more effective circulating Hb level, with fewer transfusions than previously.

PREVENTION

Population Surveys and Genetic Counseling

The availability of relatively easy and inexpensive tests, such as electronic red cell counting and MCV measurement (11), permits the screening of populations at risk and the detection of asymptomatic, previously unknown heterozygotes. Such screening programs, however, if they are to be effective, must be associated with active educational and counseling services to advise the newly diagnosed heterozygotes. It is not certain, however, that such programs will in fact decrease the incidence of marriages and pregnancies at risk of producing homozygous infants.

Prenatal Diagnosis

It is possible to detect low levels of β chain synthesis in fetal blood during the first trimester. Furthermore, techniques are being perfected for the safe aspiration of fetal red blood cells in utero (23, 24). In fact, the diagnoses of heterozygous α- and β-thalassemia have been made with first trimester fetal red cells obtained either by amnioscopy (25), or after therapeutic abortion (26). As these techniques are perfected, it may soon be possible to make definitive prenatal diagnoses of homozygous α- and β-thalassemia in pregnancies at risk.

CONCLUSIONS

Although a great deal has been learned in recent years about the molecular mechanisms involved in thalassemia, there is still no prospect, in the near future, of any molecular cure or control of the disease by gene therapy or manipulation. However, methods of diagnosis and treatment have improved and one can offer most thalassemic children nearly two decades of comfortable and useful life. Continued advances in the fields of transfusion therapy and pharmacology may progressively increase the average longevity of the patient with homozygous β-thalassemia.

ACKNOWLEDGMENTS

Some of the work described was supported in part by grants from the National Institutes of Health, Bethesda, Maryland. Dr. Forget is the recipient of a USPHS Research Career Development Award.

Literature Cited

1. Weatherall, D. G., Clegg, J. B. 1972. *The Thalassaemia Syndromes.* Oxford: Blackwell. 2nd ed. 374 pp.
2. Forget, B. G. 1974. The molecular basis of thalassemia. *Crit. Rev. Biochem.* 2: 311–42
3. Nathan, D. G., Gunn, R. B. 1966. Thalassemia: The consequences of unbalanced hemoglobin synthesis. *Am. J. Med.* 41:815–30
4. Nienhuis, A. W., Anderson, W. F. 1971. Isolation and translation of hemoglobin messenger RNA from thalassemia, sickle cell anemia and normal human reticulocytes. *J. Clin. Invest.* 50:2458–60
5. Benz, E. J. Jr., Forget, B. G. 1971. Defect in messenger RNA for human hemoglobin synthesis in beta thalassemia. *J. Clin. Invest.* 50:2755–60
6. Housman, D., Forget, B. G., Skoultchi, A., Benz, E. J. Jr. 1973. Quantitative deficiency of chain specific globin messenger ribonucleic acids in the thalassemia syndromes. *Proc. Nat. Acad. Sci. USA* 70:1809–13
7. Kacian, D. L. et al 1973. Decreased globin messenger RNA in thalassemia detected by molecular hybridization. *Proc. Nat. Acad. Sci. USA* 70:1886–90
8. Forget, B. G. et al 1974. Molecular genetics of hemoglobinopathies: Study by fingerprint analysis of normal and thalassemic globin messenger RNA. *J. Clin. Invest.* 53:23a (Abstr.)
9. Taylor, J. M. et al 1974. Genetic lesion in homozygous α thalassaemia (Hydrops foetalis). *Nature* 251:392–93
10. Ottolenghi, S. et al 1974. The severe form of α thalassaemia is caused by a hemoglobin gene deletion. *Nature* 251:389–92
11. Pearson, H. A., O'Brien, R. T., McIntosh, S. 1973. Screening for thalassemia trait by electronic measurement of mean corpuscular volume. *N. Engl. J. Med.* 288:351–53
12. England, J. M., Fraser, P. M. 1973. Differentiation of iron deficiency from thalassemia trait by routine blood count. *Lancet* 1:449–52
13. Mentzer, W. C. Jr. 1973. Differentiation of iron deficiency from thalassemia trait. *Lancet* 1:882
14. Torlontano, G., Tata, A., Camagna, A. 1972. A rapid screening test for thalassaemia trait. *Acta Haematol.* 48:234–38
15. Wasi, P., Disthasongchan, P., Na-Nakorn, S. 1968. The effect of iron deficiency on the levels of Hb A_2 and E. *J. Lab. Clin. Med.* 71:85–91
16. Wolman, I. J., Ortolani, M. 1969. Some clinical features of Cooley's anemia patients as related to transfusion schedules. *Ann. N.Y. Acad. Sci.* 165:407–14
17. Necheles, T. F., Chung, S., Sabbah, R., Whitten, D. 1974. Intensive transfusion program in thalassemia major: An 8 year follow-up. *Ann. N.Y. Acad. Sci.* 232:179–85
18. Piomelli, S. et al 1974. Hypertransfusion regimen in patients with Cooley's anemia. *Ann. N.Y. Acad. Sci.* 232: 186–92
19. O'Brien, R. T. 1974. Ascorbic acid enhancement of desferrioxamine-induced

urinary iron excretion in thalassemia major. *Ann. N.Y. Acad. Sci.* 232:221–25
20. Modell, C. B., Beck, J. 1974. Long-term desferrioxamine therapy in thalassemia. *Ann. N.Y. Acad. Sci.* 232:201–10
21. Barry, M., Flynn, D. M., Letsky, E. A., Risdon, R. A. 1974. Long-term chelation therapy in thalassemia major: Effect on liver iron concentration, liver histology and clinical progress. *Brit. Med. J.* 2:16–20
22. Graziano, J. H., Grady, R. W., Cerami, A. 1974. The identification of 2,3-dihydroxybenzoic acid as a potentially useful iron-chelating drug. *J. Pharmacol. Exp. Ther.* 190:570–75
23. Kan, Y. W., Valenti, C., Guidotti, R., Carnazza, V., Rieder, R. F. 1974. Fetal blood-sampling in utero. *Lancet* 1: 79–80
24. Hobbins, J. C., Mahoney, M. J. 1974. In utero diagnosis of hemoglobinopathies. Technic for obtaining fetal blood. *N. Engl. J. Med.* 290:1065–67
25. Kan, Y. W., Bellevue, R., Rieder, R., Valenti, C., Dozy, A. M. 1974. Prenatal diagnosis of α thalassemia. *Clin. Res.* 22:374A (Abstr.)
26. Chang, H., Modell, C. B., Huehns, E. R., Nathan, D. G. 1974. Antenatal diagnosis of the β-thalassemia gene. *Pediat. Res.* 8:398 (Abstr.)
27. Forget, B. G., Housman, D., Benz, E. J. Jr. 1974. Separation of human α and β globin messenger RNAs. *Blood* (Abstr.) 44:905

ENDOCRINE THERAPY OF BREAST CANCER

❖7157

William Leo McGuire, M.D.
Department of Medicine, University of Texas Health Science Center, San Antonio, Texas 78284

INTRODUCTION

The first demonstration that the growth of a human cancer could be dramatically influenced by endocrine secretions was made by Beatson, who in 1896 observed the regression of metastatic breast cancer lesions following ovariectomy (1). More than 75 years later, the biochemical mechanism of the phenomenon remains incompletely understood. Part of the problem in unraveling this mechanism is that either ovariectomy, adrenalectomy, or hypophysectomy may result in objective tumor regression (2); experiments to discover which hormones are most critical have not entirely resolved the question. In view of this uncertainty about the roles of particular hormones, 30% of human tumors which respond to any endocrine therapy have thus far been labeled simply "hormone-dependent."

Recent advances in methodology, particularly in radioimmunoassay, now permit us to define almost completely the hormonal environment of a growing mammary tumor. Advances in subcellular biochemistry are similarly permitting us to understand how a cell is recognized as a target cell by a particular hormone in the environment. Specific cytoplasmic protein receptors for steroid hormones and surface membrane receptors for polypeptide hormones have been identified. This has turned out to be significant for breast cancer research because of the observation that both human and experimental breast tumors which regress following endocrine therapy usually contain the specific cytoplasmic receptor for estrogen, while nonresponsive (autonomous) tumors are often deficient in estrogen receptor (3).

These findings, which will be discussed in detail, have led to the following concepts:

1. Normal mammary cells contain cytoplasmic or membrane receptor sites for

2. When malignant transformation occurs, the cell may retain all or only part of the normal population of receptor sites. If the cell retains the receptor sites, its growth and function is potentially capable of being regulated by the hormonal environment as in a normal cell; however, if the receptors are lost from the cell as a consequence of malignant transformation, the cell is no longer recognized as a target cell by circulating hormones and endocrine control is absent.

3. The absence of specific receptors in mammary tumor tissue may therefore indicate hormonal autonomy, thus aiding in selecting the 30% of breast cancer patients who will benefit from endocrine therapy.

The primary purpose of this review is to examine the role of estrogen receptors, the only receptors which have been studied to date with respect to all three concepts above, in light of recent advances in endocrinology, especially in the biochemistry of estrogen action. Wherever possible, data derived from human breast tumors will be cited, but most of the studies in hormone-dependent breast carcinoma are derived from carcinogen-induced rat mammary tumors; the methods of induction and ablation and the endocrine aspects of these tumors have been extensively reviewed (4). Finally, the correlation of estrogen receptor levels with response to endocrine therapy and other clinical data will be examined in 380 collected cases from several centers.

ESTROGEN RECEPTORS

Biological Effect of Estrogen on Breast Tumors

Estrogen has been shown to act directly on the normal mammary gland to promote growth and differentiation. However, estrogen also stimulates the release of pituitary prolactin, and prolactin also acts upon the mammary cell. Since estrogen cannot support mammary tumor growth in the absence of a pituitary, whereas prolactin reportedly supports normal mammary gland and mammary tumor growth in the absence of ovaries and adrenals, estrogen is considered by many to play only a secondary role in breast tumor growth and regression (5). In this regard, however, it may be significant that experiments showing prolactin stimulation of tumor growth in the absence of ovarian steroids were of brief duration; if dimethylbenzanthracene (DMBA) tumor-bearing rats are ovariectomized and simultaneous lesions are placed in the median eminence to increase prolactin release, these tumors grow at an accelerated pace for only 10 to 12 days and then regress, even though prolactin levels remain elevated (6). Furthermore, it has been reported that the transplantation survival of the MTW9 rat mammary tumor requires ovarian hormones, and MTW9 tumors transplanted to rats immunized with estradiol-bovine serum albumin (BSA) conjugates will grow less well than in untreated controls (7). One might summarize the evidence bearing on the role of physiological estrogen levels as follows: estrogens are probably essential but not sufficient for growth of certain mammary tumors.

Another important effect of estrogens is the regression of mammary tumors following pharmacologic doses. This seemingly paradoxical effect of estrogen appears to act by interfering with the prolactin stimulation of growth, since the effect can be overcome by increasing endogenous or exogenous prolactin (8).

Localization of Estrogens in Responsive Tumors

In 1959, two laboratories reported that radioactively labeled estrogen injected in vivo into experimental animals was localized in those organs which respond to estrogen or excrete it (9, 10). Soon after, breast cancer patients scheduled for adrenalectomy to remove the source of circulating estrogens were given tritiated hexestrol just prior to surgery. It was discovered that the tumor metastases of patients responding to the adrenalectomy concentrated a larger fraction of ^{3}H-hexestrol than those of patients who failed to respond (11), as if only responsive tumors behaved as estrogen target tissues.

Parallel studies in rats bearing mammary carcinomas induced by DMBA achieved similar results (12), and hormone-dependent tumors were also found to take up more estrogen in vitro than autonomous tumors (13). This in vitro uptake could be completely inhibited by synthetic estrogen analogues while the relatively low uptake by other tissues such as muscle could not be inhibited, indicating specificity of the uptake into tumors (14, 15). From these results, Jensen proposed that the in vitro technique might be extended to human tumor tissue samples to predict the response to adrenalectomy. By this time, however, estrogen receptor had been discovered in target tissues including tumors (12) and appeared to be responsible for the specific uptake of estrogen by these tissues. Direct studies of the presence and role of receptor in mammary tumors followed, and included the possibility of using the presence of the receptor to predict hormone dependence. Our own part in these studies is described in the following four sections.

Hormone-Dependent Rat Mammary Tumors

We first demonstrated estrogen receptor (ER) in the cytoplasm of hormone-dependent tumors by incubating the cytosol fraction with tritiated estradiol (^{3}H-E) and applying it to a Sephadex G-100 column. A majority of the ^{3}H-E was eluted bound to the macromolecular fraction. Sucrose gradient centrifugation of these cytosols usually revealed two peaks of protein-bound radioactivity—one at 8 S and another at 4 S. Whereas the 8 S binding peak always represents specific receptor-estrogen (R-E) interaction, the 4 S peak may contain both specific and nonspecific binding components. In 0.3 M KCl, specifically bound ^{3}H-E in either uterus or tumor cytosol migrates exclusively at 4 S, suggesting that the 8 S binding molecules dissociate into subunits at high ionic strength. Physiological ionic strength (0.15 M salt) causes the receptor to sediment at 6 S. Although the 8 S species obtained in low salt gradients is therefore probably a nonphysiological form, it is a fortuitous one where separating specific receptor is concerned since all proteins which bind estradiol nonspecifically sediment at 4.6 S or less. The significance of these sedimentation forms is still in question, especially in the light of the finding that polyanions such as heparin can cause receptor to sediment at any rate between 8 S and 4 S, depending on the polyanion concentration (16).

Evaluation of the affinity of receptor for estrogen requires a method for separating bound from free estradiol. The dextran-coated charcoal method as recently modified (17) has proved both accurate and convenient for this purpose. We find a single class of high affinity binding sites with a binding constant (Kd) of the order of 2×10^{-10} M in hormone-dependent DMBA-induced tumors. The binding of ^{3}H-E is estrogen-specific, since it is inhibited by low concentrations of unlabeled estrogen but not by 1000- to 10,000-fold excesses of hydrocortisone, progesterone, or testosterone (18).

Clearly, then, hormone-dependent rat mammary tumors possess cytoplasmic estrogen receptor, and the properties of this receptor are essentially those described for other target tissues. And like other target tissues, these tumors accumulate injected estrogen in their cell nuclei, apparently still bound to a 4–5 S form of the receptor. We will consider the nuclear binding of estrogen in more detail later.

Autonomous Rat Mammary Carcinoma

In this laboratory we have used the R3230AC transplantable mammary carcinoma described in detail by Hilf et al as one example of a breast tumor that does not regress after ovariectomy (19).

We first found that nuclei from this tumor do not appreciably concentrate ^{3}H-E injected in vivo (20). This was attributed to the very low level of cytoplasmic receptor after we demonstrated a tenfold lower estrogen binding capacity in R3230AC cytosol than in a representative hormone-dependent DMBA tumor cytosol. The affinity constants were equal. The R3230AC 8 S receptor as revealed by sucrose gradient sedimentation was also very low (21).

We also considered the possibility that R3230AC might have lost the ability to bind R-E to chromatin, providing another reason for its autonomy. We therefore prepared chromatin from purified R3230AC nuclei and measured the ability of cytosols containing various amounts of receptor to bind estrogen to chromatin. We found that cytoplasm from R3230AC tumor, muscle, or brain failed to bind estrogen to the tumor chromatin. This failure is due to the paucity or total lack of receptor in these tissues. However, the complex of estrogen with rat uterine cytosol which contains abundant receptor demonstrated remarkable binding to R3230AC chromatin. These results indicate that chromatin of these autonomous breast tumor cells possesses the capacity for extensive interaction with R-E. Consequently, the failure of R3230AC nuclei to accumulate estrogen in vivo can be localized to a deficiency of receptor in the cytoplasm (22, 23).

Nuclear Acceptor Activity

The finding that chromatin from the autonomous R3230AC tumor binds R-E led to an investigation of R-E binding by nuclei of other tissues (24). As indicated above, only target cells for estrogen possess the specific cytoplasmic R, and there have been suggestions that nuclear acceptor sites for R-E are likewise found only in target cells. In questioning these latter suggestions, we determined quantitative R-E binding to nuclei prepared from several different tissues, and also examined the recovered receptor qualitatively by sucrose sedimentation.

We found that cytoplasmic R-E bound equally well to nuclei prepared from target and nontarget tissues and furthermore that nuclear binding in vitro was strictly proportional to the available R-E with no evidence of saturation of a special class of binding sites even up to many times the level present in vivo (25). Experiments in vivo have led to similar conclusions (26).

It seems likely, then, that some mechanism other than binding to a small number of specific sites must be found to explain the action of R-E in target cell and tumor nuclei. In this regard, it is important to recall that binding of R-E in the nucleus has not yet been proved to be directly involved in mediating all responses to the hormone. It may be possible that other activities of R-E in the nucleus or even in the cytoplasm, undetected by present techniques, have major roles in determining the response.

Estrogen Receptor in Human Breast Tumors

The properties of the estrogen receptor as determined in induced hormone-dependent rat tumors have now been found in human mammary tumor cytosols as well (27). Two of these properties are employed in our laboratory to quantitate ER in human breast cancer specimens obtained at surgery (28). The first is the high affinity binding of ^{3}H-E, evaluated by equilibrating cytosol with various low concentrations of labeled hormone and then removing the unbound with dextran-coated charcoal as described earlier. Scatchard plots of the binding data reveal the receptor, if present, as a very high affinity binding component (Kd $< 1 \times 10^{-10}$ M), and permit direct extrapolation to determine the amount of this component. The second property, the sedimentation of receptor primarily at 8 S in low salt sucrose gradients, is employed to confirm the above determination by an independent method. Because part of the 4 S binding peak may also be due to specific receptor, a parallel gradient is always run with a 100-fold excess of unlabeled estradiol to demonstrate the nonspecific binding components alone.

We are now in a position to explore Jensen's original suggestion that the presence of ER in a human breast tumor might indicate that the tumor was hormone-dependent and could be made to regress by appropriate endocrine manipulation. To date this laboratory has assayed ER in 154 primary tumors and 72 metastatic tumors from surgery, for correlation with response to endocrine therapy (29).

In Figure 1 we see that the values in primary tumors range from 0 to almost 1000 mol per mg of cytosol protein. (The level of sensitivity in the two methods is such that a value of less than 3 is essentially equivalent to 0 and is considered a negative assay.) Positive ER values (>3) are found in 70% of primary specimens and 58% of metastatic specimens. We have previously speculated that the wide range of values apparent in our results is due to a combination of factors, including: (*a*) variations in epithelial vs stromal content of the tumor, (*b*) the degree of dedifferentiation of the tumor, and (*c*) the amount of endogenous estradiol secreted by the patient (since endogenous estradiol would occupy ER sites and make them unavailable for assay) (20). This last point may at least partially explain why the highest values for tumor ER are seen in postmenopausal patients.

CLINICAL CORRELATION

A number of other laboratories have likewise assayed ER in breast tumor specimens using a variety of techniques. Data on clinical response to endocrine therapy are now available in many of these cases. On July 18–19, 1974, an international workshop

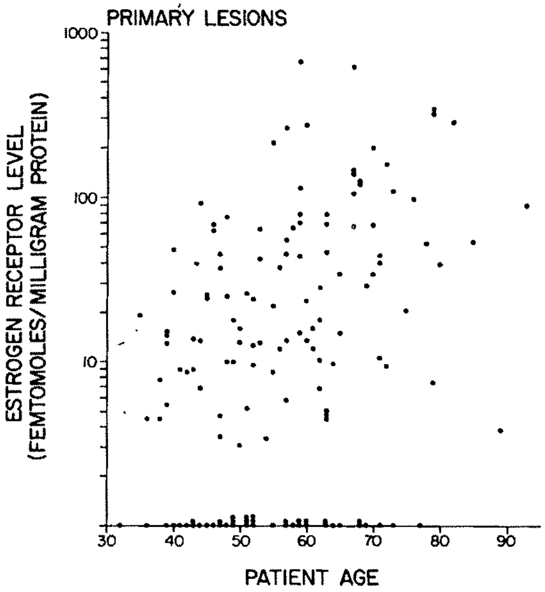


Figure 1 ER values in human primary breast specimens listed according to patient age.

was held in Bethesda, Maryland for the correlation of these data. Details of both ER assay procedures and clinical evaluation criteria were examined, and 436 treatment trials in 380 patients were ultimately accepted. I will now provide a brief overview of the data presented at the meeting, indicating the current status of ER assays in predicting response to endocrine therapies in patients with metastatic breast cancer. For details, the reader should consult the specific manuscripts published in the conference proceedings (30).

Endocrine Responsiveness

Since the organizing committee for this conference felt that clinical response data for endocrine responsiveness were as critical to the correlations as the assay data, it was arranged that participating institutions could request an extramural review of their case material. Prior to the conference, eight institutions were visited by two oncologists and a total of 531 treatment trials in 453 patients were reviewed by the following criteria:

Objective remission was defined as a 50% decrease in size of at least 50% of the measurable lesions, while other lesions are unchanged and there is no observable evidence of any new lesions. As the standard basis for determining the size of a lesion, the product of the two longest perpendicular diameters of the lesion was used. In osteolytic metastasis, evidence of healing upon roentgenography was necessary, again without increase in size or number of destructive lesions. Osteoblastic metastasis as the only measurable lesion was not an acceptable criterion for the study. For instances of multiple skin lesions, which often are impossible to measure with accuracy, photographs served as the means for evaluation of response. Neither the healing of an ulcerating lesion nor the clearing of pleural effusion was accepted as objective evidence of remission.

Using the above criteria, 436 treatment trials in 380 patients were evaluated for both clinical responsiveness to endocrine therapy and ER assay data. These data are summarized in Table 1.

ABLATIVE THERAPY Thirty three percent of 211 treatment trials yielded objective tumor regressions. Of the 94 trials in patients with negative tumor ER values, only 8 (8%) were successful, whereas 59 (55%) of the 107 trials in patients with positive tumor ER values succeeded. Patients with borderline tumor ER values had a 30% response rate.

ADDITIVE THERAPY Thirty four percent of 170 trials yielded objective tumor regressions. Of the 82 trials in patients with negative tumor ER values, 7 (8%) were successful, whereas 51 (60%) of the 85 trials in patients with positive tumor ER values succeeded.

MISCELLANEOUS THERAPY Twenty seven percent of 55 trials yielded responses to a variety of endocrine therapies including antiestrogens, aminoglutethimide, etc. Of 32 trials in patients with negative tumor ER values, 5 (16%) were successful, whereas 10 (43%) of 23 trials in patients with positive ER values succeeded.

General Considerations

There was also strong but not necessarily conclusive evidence regarding the following points:

Table 1 Objective breast tumor regressions according to ER assay and type of therapy as judged by extramural review

Therapy	ER+	ER–	ER±
Adrenalectomy	32/66	4/33	3/8
Castration	25/33	4/53	0/2
Hypophysectomy	2/8	0/8	—
Total:	59/107 = 55%	8/94 = 8%	3/10 = 30%
Androgen	12/26	2/24	0/1
Estrogen	37/57	5/58	0/2
Glucocorticoid	2/2	—	—
Total:	51/85 = 60%	7/82 = 8%	0/3 = 0%
Antiestrogens	8/20	5/27	—
Other	2/3	0/5	—
Total:	10/23 = 43%	5/32 = 16%	—

1. Most investigators agreed that tumors from postmenopausal patients contained higher ER values than those from premenopausal patients. This was thought due at least in part to endogenous estrogens in premenopausal patients occupying receptor sites, thus making them unavailable for assay.
2. There was no correlation between the histological type of tumor and the presence of ER. However, there was a suggestion that morphologically undifferentiated tumors (grade III) less often contained ER.
3. There was a negative correlation between the size of the primary tumor and the presence of ER. It was speculated that perhaps the larger tumors were more necrotic and thus contained less ER.
4. There was no correlation between the clinical stage of the patient and tumor ER.
5. There was no correlation between positive or negative axillary lymph nodes and the presence of ER in the primary tumor.
6. There was no correlation with the location of the tumor in the breast and the presence of ER.
7. The percentage of primary tumors reported to contain ER has risen from ∿50% a few years ago to 70–85% in several recent series. This rise has been attributed to increased sensitivity of the methods employed.
8. There is no conclusive correlation between disease-free interval and ER.
9. There is no correlation between the absolute ER value and the duration of remission.
10. There does not appear to be a linear relationship between the absolute value for tumor ER and the percentage of patients responding to endocrine therapy. In certain series patients with borderline ER values are able to respond with nearly the same frequency as those patients with high ER values.
11. The prognostic significance of ER values for response to endocrine therapies does not seem to depend on the type of tissue that is biopsied for the assay. ER values from primary tumors or metastatic tumors predict equally well.
12. There were little data on male breast cancer or on possible ethnic variations.
13. There were few data presented on the constancy of ER values in the primary and subsequent metastases. More data are needed on sequential biopsies for ER.

COMMENTARY

The data presented leave little doubt that estrogen receptor assays can be helpful to predict the results of endocrine therapy for metastatic breast cancer. It is clear that if a patient has a negative tumor ER value, the chances of tumor regression in response to endocrine therapy are minimal. It seems that a large number of patients can thus be spared unrewarding major endocrine ablative therapy if ER assays are performed routinely. If the tumor ER value is positive, the response to endocrine therapy is 55–60%. This single piece of evidence when coupled with available clinical prognostic factors such as menopausal status, disease-free interval, site of dominant lesion, and especially response to previous hormonal therapies should permit the practicing oncologist to select or reject endocrine therapy with considerable confidence.

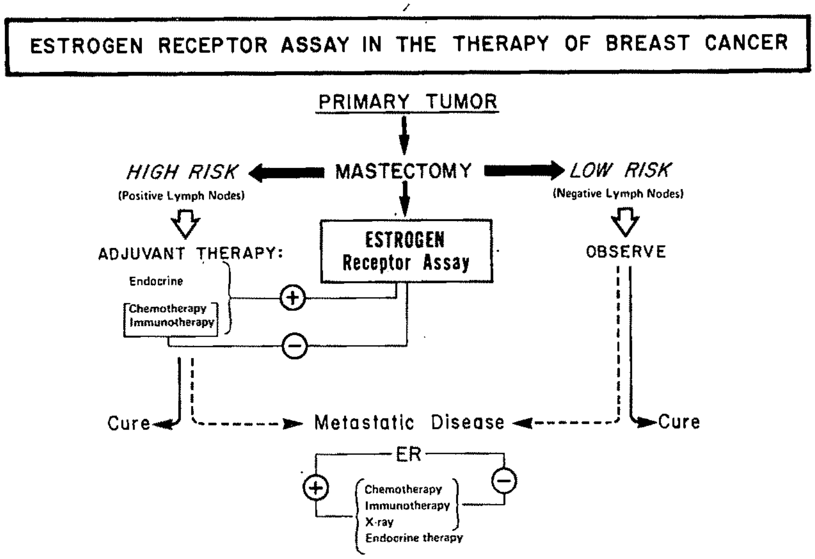


Figure 2 A guide for the use of tumor ER values in selecting endocrine therapy for patients with breast cancer.

It is clear that at the time of original mastectomy, patients can be divided into high and low risk groups depending upon the status of their axillary nodes. There is a great deal of investigation now under way to determine if adjuvant therapy might improve the prognosis in the high risk group (Figure 2). Estrogen receptor assays at the time of original surgery could be used to determine whether an endocrine therapy should be a part of the adjuvant regimen. For example, one could choose to include pharmacological doses of androgen or estrogen. In dealing with overt metastatic disease, the same rationale should be employed. If the primary tumor or a biopsy of metastatic tissue reveals ER, then an endocrine therapy should be considered as part of the overall therapeutic scheme.

Finally, the ER data should be considered in light of current theories of hormone dependence in breast cancer. When malignant transformation occurs, the cell may retain all or only part of the normal population of receptor sites. If the cell retains the receptor sites, its growth and function is potentially capable of being regulated by the hormonal environment as in a normal cell; however, if the receptors are lost from the cell as a consequence of malignant transformation, the cell is no longer recognized as a target cell by circulating hormones and endocrine control is absent. This could explain why patients fail to respond to endocrine therapy if their tumors lack ER.

Why then did not all of the patients with positive tumor ER values respond to endocrine therapy? It should be emphasized that ER is only a part of the complex hormonal control system known to regulate mammary tumor growth. The biochemical mechanism by which prolactin recognizes a breast tumor as a target cell must be equally important (31), and other hormones are undoubtedly involved as well. Nevertheless, the ER data suggest that loss of endocrine control is closely related to the loss of ER. Does this mean that receptors for other hormones are lost in conjunction with ER and will also be absent in the autonomous mammary tumor cell? Or does ER somehow occupy a unique position in determining hormone dependence? Until we have the answers to these questions it may be useful to consider simply that the absence of ER in a tumor cell is one indication of the relative departure from normality that has occurred in the process of malignant transformation, and may therefore be useful in predicting the loss of hormone dependence. Because other biochemical events involving estrogen itself as well as prolactin and other hormones are probably prerequisites for complete endocrine regulation, we can predict that other biochemical lesions have occurred in patients in whom endocrine therapy failed despite positive tumor ER levels. We still have a great deal to learn about the subcellular biochemistry of hormone action in breast cancer tissue, and future investigation will undoubtedly lead to further improvements in therapy for patients with breast cancer.

ACKNOWLEDGMENTS

The clinical workshop data were contributed by the following participating institutions: Alton Ochsner Medical Foundation, New Orleans, Louisiana; Ben May Laboratory for Cancer Research, Chicago, Illinois; Case Western Reserve University, Cleveland, Ohio; Institut Jules Bordet, Brussels, Belgium; Michigan Cancer Foundation, Detroit, Michigan; Middlesex Hospital & Medical School, London, England; Netherlands Cancer Institute, Amsterdam, The Netherlands; Norsk Hyro's Institute for Cancer Research, Oslo, Norway; St. Mary's Hospital Medical School, London, England; The Imperial Cancer Research Fund, London, England; University of Hamburg, Hamburg, Germany; University of Oregon Medical School, Portland, Oregon; University of Rochester School of Medicine and Dentistry, Rochester, New York; University of Texas Health Science Center, San Antonio, Texas; University of Uppsala, Uppsala, Sweden; University of Western Australia, Perth, Western Australia. These studies were supported in part by U.S. Public Health Service Grants CA-11378 and CB-23862, and American Cancer Society Grant BC-23C.

Literature Cited

1. Beatson, G. T. 1896. On the treatment of inoperable cases of carcinoma of the mamma: Suggestions for a new method of treatment, with illustrative cases. *Lancet* 2:104–7, 162–65
2. Dao, T. L. 1972. Ablation therapy for hormone-dependent tumors. *Ann. Rev. Med.* 23:1–18
3. McGuire, W. L., Chamness, G. C., Costlow, M. E., Shepherd, R. E. 1974. Hormone dependence in breast cancer. *Metabolism* 23:75–100
4. Dao, T. L. 1964. Carcinogenesis of mammary gland in rat. *Progr. Exp. Tumor Res.* 5:157–216
5. Pearson, O. H., Llerena, O., Llerena, L., Molina, A., Butler, T. 1969. Prolactin-dependent rat mammary cancer: A model for man? *Trans. Assoc. Am. Physicians* 82:225–37
6. Sinha, D., Cooper, D., Dao, T. L. 1973. The nature of estrogen and prolactin effect on mammary tumorigenesis. *Cancer Res.* 33:411–14

7. Caldwell, B. V., Tillson, S. A., Esber, H., Thorneycroft, I. H. 1971. Survival of tumors after immunization against oestrogens. *Nature (London)* 231: 118–19
8. Meites, J., Cassell, E., Clark, J. 1971. Estrogen inhibition of mammary tumor growth in rats: Counteraction by prolactin. *Proc. Soc. Exp. Biol. Med.* 137:1225–27
9. Glascock, R. F., Hoekstra, W. G. 1959. Selective accumulation of tritium-labelled hexoestrol by the reproductive organs of immature female goats and sheep. *Biochem. J.* 72:673–82
10. Jensen, E. V., Jacobson, H. I. 1960. *Biological Activities of Steroids in Relation to Cancer,* ed. G. Pincus, E. P. Vollmer, 161–74. New York: Academic
11. Folca, P. J., Glascock, R. F., Irvine, W. T. 1961. Studies with tritium-labelled hexoestrol in advanced breast cancer. *Lancet* 2:796–98
12. King, R. J. B., Gordon, J., Cowan, D. M., Inman, D. R. 1966. The intranuclear localization of (6, 7-^{3}H) -oestradiol-17β in dimethylbenzanthracene-induced rat mammary adenocarcinoma and other tissues. *J. Endocrinol.* 36:139–50
13. Mobbs, B. G. 1966. The uptake of tritiated oestradiol by dimethylbenzanthracene-induced mammary tumours of the rat. *J. Endocrinol.* 36:409–14
14. Jensen, E. V., DeSombre, E. R., Jungblut, P. W. 1967. *Endogenous Factors Influencing Host-Tumor Balance,* ed. R. W. Wissler, T. L. Dao, S. Wood Jr., 15–30. Chicago: Univ. Chicago
15. Terenius, L. 1968. Selective retention of estrogen isomers in estrogen-dependent breast tumors of rats demonstrated by in vitro methods. *Cancer Res.* 28:328–37
16. Chamness, G. C., McGuire, W. L. 1972. Estrogen receptor in the rat uterus. Physiological forms and artifacts. *Biochemistry* 11:2466–72
17. McGuire, W. L., De La Garza, M. 1973. Improved sensitivity in the measurement of estrogen receptor in human breast cancer. *J. Clin. Endocrinol. Metab.* 37:986–89
18. McGuire, W. L., Julian, J. A. 1971. Comparison of macromolecular binding of estradiol in hormone-dependent and hormone-independent rat mammary carcinoma. *Cancer Res.* 31: 1440–45
19. Hilf, R., Michel, I., Bell, C. 1967. Biochemical and morphological responses of normal and neoplastic mammary tissue to hormonal treatment. *Recent Progr. Horm. Res.* 23:229–90
20. McGuire, W. L., Chamness, G. C. 1973. Studies on the estrogen receptor in breast cancer. *Advan. Exp. Med. Biol.* 36:113–36
21. McGuire, W. L., Julian, J. A., Chamness, G. C. 1971. A dissociation between ovarian-dependent growth and estrogen sensitivity in mammary carcinoma. *Endocrinology* 89:969–73
22. McGuire, W. L., Huff, K., Jennings, A., Chamness, G. C. 1972. Mammary carcinoma: a specific biochemical defect in autonomous tumors. *Science* 175: 335–36
23. McGuire, W. L., Huff, K., Chamness, G. C. 1972. Temperature-dependent binding of estrogen receptor to chromatin. *Biochemistry* 11:4562–65
24. Chamness, G. C., Jennings, A. W., McGuire, W. L. 1973. Oestrogen receptor binding is not restricted to target nuclei. *Nature (London)* 241:458–60
25. Chamness, G. C., Jennings, A. W., McGuire, W. L. 1974. Estrogen receptor binding to isolated nuclei. A nonsaturable process. *Biochemistry* 13: 327–31
26. Shepherd, R. E., Huff, K., McGuire, W. L. 1974. Non-interaction between in vivo and cell free nuclear binding of estrogen receptor. *Endocrine Res. Commun.* 1:73–85
27. McGuire, W. L., De La Garza, M. 1973. Similarity of the estrogen receptor in human and rat mammary carcinoma. *J. Clin. Endocrinol. Metab.* 36:548–52
28. McGuire, W. L. 1973. Estrogen receptors in human breast cancer. *J. Clin. Invest.* 52:73–77
29. McGuire, W. L., Pearson, O. H., Segaloff, A. 1975. *Estrogen Receptor in Human Breast Cancer,* ed. W. L. McGuire, P. P. Carbone, E. P. Vollmer. New York: Raven. In press
30. McGuire, W. L., Carbone, P. P., Sears, M. E., Escher, G. C. 1975. *Estrogen Receptor in Human Breast Cancer,* ed. W. L. McGuire, P. P. Carbone, E. P. Vollmer. New York: Raven. In press
31. Costlow, M. E., Buschow, R. A., McGuire, W. L. 1974. Prolactin receptors in an estrogen receptor-deficient mammary carcinoma. *Science* 184: 85–86

Copyright 1975. All rights reserved

INTERCONVERSIONS OF APOLIPOPROTEIN FRAGMENTS

❖7158

Conrad B. Blum, M.D. and Robert I. Levy, M.D.
Molecular Disease Branch, National Heart and Lung Institute, National Institutes of Health, Bethesda, Maryland 20014

Over the past decade attention has focused on the protein moieties of the lipoproteins, the lipoprotein apoproteins. It has become clear that they hold the key to our understanding of normal as well as abnormal lipid transport. We have realized that there is a heterogeneous group of lipoprotein apoproteins that differ biologically, immunologically, and physicochemically, some with apparently only structural, others with more functional roles in normal lipoprotein metabolism.

APOPROTEIN NOMENCLATURE AND STRUCTURE

The lipoprotein particles are usually operationally classified either according to ultracentrifugal flotation or electrophoretic mobility (1). Chylomicrons, the carriers of exogenous dietary triglyceride, are the largest and lightest of the lipoprotein particles. They remain at the origin on paper electrophoresis. The very low density lipoproteins (VLDL) are the major carriers of endogenously produced triglycerides; they occur in the density range 0.93–1.006 g/ml and have pre-beta mobility on paper electrophoresis. The low density lipoproteins (LDL) are the products of VLDL catabolism. They normally account for the bulk of plasma cholesterol; they have density 1.019–1.063 and have beta mobility on paper electrophoresis. The high density lipoproteins (HDL) are importantly involved in cholesteryl ester metabolism; they occur in the density range 1.063–1.21 and have alpha electrophoretic mobility.

Heterogeneity of the apoproteins of VLDL and HDL was demonstrated immunochemically about a decade ago (2, 3). Since then several different proteins have been isolated by various chromatographic and electrophoretic techniques (4–6). Unfortunately, overlapping systems of nomenclature for these apolipoproteins have led to some confusion. The different systems are reviewed in Table 1. The A, B, C system of Alaupovic (7) is coming into increasingly wider use and will be used throughout this text. The relative distribution of the different lipoprotein apoproteins is reviewed in Table 2.

Table 1 The human plasma apolipoproteins

Apoprotein	Alternate nomenclature	Molecular weight	Biologic role
ApoA-I	R-thr, apoLp-glnI, fraction III	26,000	Activator for LCAT[a]
ApoA-II	R-gln, apoLp-glnII, fraction IV	17,380	Structural protein of HDL
ApoB	R-ser, apoLDL	?	Involved in transport of triglycerides out of cells
ApoC-I	R-val, apoLp-ser, apoD-I, fraction V	6500	?
ApoC-II	R-glu, apoLp-glu apoD-II, fraction V	10,000	Activator for LPL[b]
ApoC-III	R-ala, apoLp-ala, apoD-III, fraction V	8764	?

[a] LCAT — lecithin:cholesterol acyltransferase.
[b] LPL — lipoprotein lipase.

Table 2 Distribution of plasma apolipoproteins

Apoprotein	Chylomicrons	VLDL	LDL	HDL
	(Percentage of total lipoprotein protein)			
ApoA-I	?	T>	—	72
ApoA-II	?	?	—	24
ApoB	30	40[a]	>95	T>
ApoC-I	15	10[a]	T>	1
ApoC-II	15	10[a]	T>	1
ApoC-III	45	30[a]	T>	3

[a] Ratio of apoC/apoB varies .85–3 with VLDL subspecies, denser VLDL having relatively more apoB (17). T—trace.

ApoA-I and apoA-II are the two major peptides of HDL and account for 95–98% of its protein (8). They have molecular weights 26,000 and 17,380. ApoA-II appears to have a primarily structural role. It binds with and stabilizes the lipid in the HDL particle. ApoA-I is probably held by protein-protein interactions with apoA-II (9). ApoA-I is the activator protein for lecithin: cholesterol acyltransferase (LCAT) (10), the enzyme responsible for most cholesterol esterification in man (11). With its complement of ApoA-I, HDL is the physiological substrate for LCAT. The amino acid sequence of ApoA-II is known (12) while that of the amino-terminal 39 residues of ApoA-I has been reported (13).

ApoB seems to have a major structural role in chylomicrons, VLDL, and LDL analogous to that of apoA-II in HDL. Little is known about its structure because of its strong tendency to aggregate and precipitate when separated from lipids. Thus, the molecular weight of the basic apoB subunit has been reported over the entire range from 25,000 to 250,000 (14, 15). It is the only protein of LDL and accounts for about half of the protein mass of VLDL (16, 17).

The low molecular weight (7000–12,000) *apoC peptides* are major components of VLDL (17) and chylomicrons (18) but minor components of HDL (8). They consist of at least three different peptides, apoC-I, apoc-II, and apoC-III. There is also an apoC-IV which differs from apoC-III only by one sialic acid residue (19). The amino acid sequences of apoC-I and apoC-III and apoC-IV are known (20, 21). Their functional roles are unclear however. ApoC-II has not been sequenced. It has been shown, however, to be a coenzyme for lipoprotein lipase (LPL) (22–24). The triglyceride-rich VLDL and chylomicrons which have a high content of apoC-II are the natural substrates for this enzyme.

The different apoproteins tend to bind noncovalently to particular lipids, forming particles of specific densities. Over a decade ago Scanu et al showed that the apoproteins of HDL when incubated with plasma or lipid formed complexes having the same density as native HDL (25). They later extended this observation to isolated apoA-I and apoA-II (26). The apoC peptides bind to HDL but have a greater affinity for VLDL and triglyceride (26). A threefold relative enrichment in triglycerides as compared to other lipids can be demonstrated in the lipoprotein complexes obtained by sonicating apoC with HDL lipids (26). Eisenberg et al have shown that the larger (and more triglyceride-laden) VLDL particles contain a greater amount of apoC as compared to apoB (17).

Optical techniques have demonstrated an increase in the orderliness of structure of apoC-III (27), apoA-I (28), and apoA-II (28) when these peptides are placed in lipid emulsions. Presumably, this represents noncovalent binding between protein and lipid.

METABOLISM OF VERY LOW DENSITY LIPOPROTEINS

A precursor-product relationship between VLDL and LDL was first suggested in 1957 by the radioactive tracer studies of Gitlin et al (29). Many subsequent studies have confirmed this using isotopically labeled protein (30), triglyceride (31), or other techniques (32, 33). Sophisticated mathematical analyses (33a) indicate that VLDL apoB is probably the sole source of LDL protein in man.

Compositional considerations dictate that VLDL lose triglyceride and apoC in their unidirectional conversion to LDL. Lipoprotein lipase-rich plasma can produce in vitro conversion of VLDL to an intermediate density lipoprotein (IDL) having density 1.006–1.019 (34). The loss of VLDL-triglyceride in vivo is mediated by LPL, while that of apoC is probably secondary, resulting from the strong affinity of those peptides for triglycerides and/or HDL. In this process, HDL acts between tides of dietary triglyceride as a repository for apoC-II, the activator of LPL. Fractionation of VLDL into density classes has shown that the particles poorest in triglycerides have the lowest amount of apoC compared to apoB (17). The LDL and IDL particles produced from VLDL by LPL-rich plasma appear to have the same amount of apoB per particle as do the original VLDL (35), suggesting the conversion of one particle of VLDL to one particle of LDL.

While catabolism of VLDL results in an LDL particle rich in apoB, the apoC is transferred to HDL (36). Even in vitro there is very rapid exchange of apoC between VLDL and HDL (30). The transfer of apoC from HDL to chylomicrons has been observed with dietary fat loads (37). Evidence has also been accumulated that suggests that apoC primarily enters the plasma with HDL and transfers to glyceride-rich VLDL and chylomicrons while in the circulation (33a). The brisk lipolysis and resultant VLDL catabolism which follows intravenous injection of heparin dramatically speeds the transfer of apoC back from VLDL to HDL (37). This is evident within minutes. The transfer of apoB to an intermediate density range (1.006–1.019) is also augmented by heparin, but to a lesser extent.

METABOLISM OF THE MAJOR HIGH DENSITY LIPOPROTEIN PEPTIDES

The metabolic pathways of HDL and its major peptides have been much less thoroughly studied than have those of VLDL and LDL. Early studies were performed without knowledge of the heterogeneity of HDL peptides. They did show an apparently monoexponential decay with half-life of about five days when ^{131}I-HDL was injected into man (38, 39). Scanu & Hughes detected no difference in plasma decay after injection into mice of native or delipidated HDL (38).

Eisenberg et al studied the decay in rats of human ^{125}I-HDL peptides and reported a much shorter half-life for apoA-II (9 hr) than for apoA-I (15 hr) (40). They also noted in the rat some disparity among the decay rates of the peptides of rat ^{125}I-HDL. One interpretation of these results is that apoA-I is able to dissociate itself from the HDL particle and thus is metabolized independently of apoA-II. ApoA-I is known to bind very loosely to HDL (2, 41), probably by protein-protein interaction with apoA-II (9). In fact, pure apoA-I particles can be identified immunochemically free of apoA-II (42). It is possible that they are artifactual here, however, resulting from their loose binding and the isolation procedures.

PATHOPHYSIOLOGY

Abetalipoproteinemia

Abetalipoproteinemia is a familial recessively inherited disorder characterized clinically by malabsorption, retinitis pigmentosa, acanthocytosis, and neurologic signs (1). In this condition, apoB is totally absent from the blood. Consequently, LDL, VLDL, and chylomicrons are absent. ApoC, however, is found in the plasma of patients with abetalipoproteinemia. This experiment of nature suggests that apoB is absolutely required for the production of chylomicrons and the formation of VLDL.

Abetalipoproteinemia has been used as a model to test the procursor-product relationship between VLDL and LDL. After normal VLDL is injected into an abetalipoproteinemic patient, LDL is found in the plasma (33).

Familial Hypobetalipoproteinemia

This dominantly inherited disorder is characterized by low but definitely present levels of LDL. Radioactive tracer techniques using ^{125}I-LDL show a low synthetic rate of LDL-protein (apoB) but a normal fractional rate of apoB catabolism (43). Since adverse medical consequences rarely occur in hypobetalipoproteinemia, we can infer that apoB is normally produced in a many-fold excess of its requirement.

Tangier Disease

In this recessively inherited disorder, normal HDL is absent, although very low levels of an abnormal HDL, HDL_T are present (44). This abnormal HDL is quite deficient in apoA-I. In addition, lipoprotein particles with apoA-II as the only apoprotein have been reported (45). These are never found in the normal and they may represent HDL_T. The relationship of the protein defect in Tangier disease to clinical aspects of the disorder including the deposition of cholesteryl esters in many tissues remains obscure.

Familial Type II Hyperlipoproteinemia

This dominantly inherited disorder, whose hallmark is elevated levels of apoB, is very highly associated with coronary artery disease (46). Langer et al (47) employing ^{125}I-LDL showed that the defect in type II is primarily in clearance or removal of apoB (LDL) from the circulation; the production rate of apoB was normal. This observation has recently been extended by Brown and Goldstein (48) who presented strong evidence that defective binding of LDL to cell membrane receptor sites is at the root of the low fractional catabolic rate. A binding defect was demonstrated in fibroblasts cultured from patients with the homozygous form of type II. They have furthermore demonstrated a similar but less severe defect in binding of LDL in heterozygous patients similar to those studied by Langer et al (49).

The Familial Hypertriglyceridemias

Hyperlipoproteinemias type I (elevated chylomicrons), type IV (elevated VLDL), and type V (elevated chylomicrons and VLDL) all have hypertriglyceridemia in common (50). Type I is further distinguished by very low levels of lipolytic activity in post-heparin plasma. Mathematical modeling of the disappearance of ^{125}I-VLDL from plasma and the appearance of ^{125}I-LDL after injection of labeled VLDL yields a common kinetic description for all three disorders (33a); the delipidation rate is slow. This is estimated as the time for a newly produced VLDL particle to be converted to a particle of intermediate density (1.006–1.019). It is especially slow in type I, the most severe triglyceridemia, which is frequently associated with plasma triglyceride levels in excess of 10,000 mg/dl and episodes of pancreatitis.

Type III Hyperlipoproteinemia

This disorder has heretofore been characterized either by the presence of VLDL with beta migration on paper electrophoresis (51) or by a high ratio of VLDL cholesterol to total plasma triglycerides (52). It has recently been discovered to be associated with a specific qualitative abnormality identifiable by isotope tracer techniques (33a). In type III, a unique kinetically defined compartment is present in the VLDL density range, and there is a much higher than normal production rate of apoB. The abnormal compartment is found in none of the other types of hyperlipoproteinemia. The material in the abnormal compartment is produced from normal

VLDL and its quantity is directly related to the apoB production rate. In type III subjects effective dietary or drug therapy reduces the apoB production rate and the amount of abnormal material. Thus, it appears that in type III, the normal mechanism for catabolism of VLDL is saturated, and an abnormal mechanism is substituted.

SPECULATIONS

With the foregoing in mind, it is possible to speculate on future observations. For instance, the enzyme deficiency disorders lipoprotein lipase deficient chylomicronemia (familial type I hyperlipoproteinemia) and lecithin: cholesterol acyltransferase deficiency may be found in some cases to result from lack of the apoC-II or apoA-I activator proteins. Therapy may in these cases be directed toward stimulating the production or inhibiting the degradation of the apoprotein activators.

Some of the quantitatively minor apoproteins like apoC-I and apoC-III may be shown to play crucial physiological roles. Furthermore, the existence of other minor apoproteins has been suggested. For example, another apoprotein especially rich in arginine and normally present in apparently trace amounts has been reported elevated in type III HLDL (53). Its physiological significance at this time however is obscure.

Certainly more information and insight can be expected as studies continue into the role of the apolipoproteins in hyperlipoproteinemia and the atherosclerotic process.

Literature Cited

1. Fredrickson, D. S., Gotto, A. M., Levy, R. I. 1972. *The Metabolic Basis of Inherited Disease,* ed. J. B. Stanbury, J. B. Wyngaarden, D. S. Fredrickson, Chap. 26, 493–530. New York: McGraw. 1778 pp.
2. Levy, R. I., Fredrickson, D. S. 1966. *J. Clin. Invest.* 45:63–77
3. Levy, R. I., Fredrickson, D. S. 1965. *J. Clin. Invest.* 44:426–41
4. Brown, W. V., Levy, R. I., Fredrickson, D. S. 1970. *J. Biol. Chem.* 244:5687–94
5. Shore, V. G., Shore, B. 1973. *Biochemistry* 12:502–7
6. Camejo, G., Suarez, F. M., Munoz, V. 1970. *Biochim. Biophys. Acta* 218: 155–66
7. Alaupovic, P. 1971. *Atherosclerosis* 13:141–46
8. Scanu, A. M., Toth, J., Edelstein, C., Koga, S., Stiller, E. 1969. *Biochemistry* 8:3309–16
9. Assmann, G., Brewer, H. B. 1974. *Proc. Nat. Acad. Sci. USA* 71:989–93
10. Fielding, C. J., Shore, V. G., Fielding, P. E. 1972. *Biochem. Biophys. Res. Commun.* 46:1493–98
11. Stokke, K. T. 1971. *Biochim. Biophys. Acta* 270:156–66
12. Brewer, H. B., Lux, S. E., Ronan, R., John, K. M. 1972. *Proc. Nat. Acad. Sci. USA* 69:1304–8
13. Shore, B., Shore, V. 1972. *Proceedings of the European Society of Atherosclerosis.* Paris: Masson and Cie
14. Pollard, H., Scanu, A. M., Taylor, E. W. 1969. *Proc. Nat. Acad. Sci. USA* 64:304–10
15. Smith, R., Dawson, J. R., Tanford, C. 1972. *J. Biol. Chem.* 247:3376–81
16. Gotto, A. M., Brown, W. V., Levy, R. I., Birnbaumer, M. E., Fredrickson, D. S. 1972. *J. Clin. Invest.* 51:1486–94
17. Eisenberg, S., Bilheimer, D., Lindgren, F., Levy, R. I. 1972. *Biochim. Biophys. Acta* 260:329–33
18. Schonfeld, G., Felski, C., Manning, C. 1973. *Circulation* 48(4):112 (Abstr.)
19. Brown, W. V., Levy, R. I., Fredrickson, D. S. 1970. *J. Biol. Chem.* 245:6588–94
20. Shulman, R. et al 1972. *Circulation* 46(2):246 (Abstr.)
21. Brewer, H. B., Shulman, R., Herbert,

P., Ronan, R., Wehrly, K. 1972. *Advan. Exp. Med. Biol.* 5:280–99

22. LaRosa, J. C., Levy, R. I., Herbert, P., Lux, S. E., Fredrickson, D. S. 1970. *Biochem. Biophys. Res. Commun.* 41: 57–62
23. Havel, R. J., Shore, V. G., Shore, B., Bier, D. M. 1970. *Circ. Res.* 27:595–600
24. Krauss, R. M., Herbert, P. N., Levy, R. I., Fredrickson, D. S. 1973. *Circ. Res.* 33:403–11
25. Scanu, A., Hughes, W. L. 1960. *J. Biol. Chem.* 235:2876–83
26. Scanu, A. et al 1970. *Biochemistry* 9:1327–35
27. Hoff, H. F., Morrisett, J. D., Gotto, A. M. 1973. *Biochim. Biophys. Acta* 296:653–60
28. Lux, S. E., Hirz, R., Shrager, R. I., Gotto, A. M. 1972. *J. Biol. Chem.* 247:2598–606
29. Gitlin, D. et al 1958. *J. Clin. Invest.* 37:172–84
30. Bilheimer, D. W., Eisenberg, S., Levy, R. I. 1972. *Biochim. Biophys. Acta* 260:212–21
31. Havel, R. J., Felts, J. M., Van Duyne, C. M. 1962. *J. Lipid Res.* 3:297–308
32. Gulbrandsen, C. L., Wilson, R. B., Lees, R. S. 1971. *Circulation* 44(2):10 (Abstr.)
33. Gulbrandsen, C., Evans, V., Nichols, A. Lees, R. 1972. *Circulation* 46(2):18 (Abstr.)

33a. Hall, M. et al 1974. *Circulation* 50(3): 114 (Abstr.)

34. Shore, B., Shore, V. 1962. *J. Atheroscler. Res.* 2:104–14
35. Eisenberg, S., Bilheimer, D. W., Levy, R. I., Lindgren, F. T. 1973. *Biochim. Biophys. Acta* 326:361–77
36. Eisenberg, S., Bilheimer, D. W., Levy, R. I. 1972. *Biochim. Biophys. Acta* 280:94–104
37. Havel, R. J., Kane, J. P., Kashyap, M. L. 1973. *J. Clin. Invest.* 52:32–38
38. Scanu, A., Hughes, W. L. 1962. *J. Clin. Invest.* 41:1681–89
39. Furman, R. H., Sanbar, S. S., Alaupovic, P., Bradford, R. H., Howard, R. P. 1964. *J. Lab. Clin. Med.* 63:193–204
40. Eisenberg, S., Windmueller, H. G., Levy, R. I. 1973. *J. Lipid Res.* 14:446–58
41. Nichols, A. V., Lux, S., Forte, T., Gong, E., Levy, R. I. 1972. *Biochim. Biophys. Acta* 270:132–48
42. Albers, J., Aladjem, F. 1971. *Biochemistry* 10:3436–42
43. Levy, R. I., Langer, T., Gotto, A. M. Fredrickson, D. S. 1970. *Clin. Res.* 18:539 (Abstr.)
44. Lux, S. E., Levy, R. I., Gotto, A. M., Fredrickson, D. S. 197. *J. Clin. Invest.* 51:2505–19
45. Assmann, G., Fredrickson, D., Herbert, P., Forte, T., Heinen, R. 1974. *Circulation* 50(3):259 (Abstr.)
46. Stone, N. J., Levy, R. I., Fredrickson, D. S., Verter, J. 1974. *Circulation* 49: 476–88
47. Langer, T., Strober, W., Levy, R. I. 1972. *J. Clin. Invest.* 51:1528–36
48. Brown, M. S., Goldstein, J. L. 1974. *Proc. Nat. Acad. Sci. USA* 71:788–92
49. Goldstein, J. M., Brown, M. S. 1974. *Trans. Ass. Am. Phys.*, Vol. 87. In press
50. Fredrickson, D. S., Levy, R. I. 1972. See Ref. 1, Chap. 28, pp. 545–614
51. Fredrickson, D. S., Levy, R. I., Lindgren. 1968. *J. Clin. Invest.* 47:2446–57
52. Hazzard, W. R., Porte, D., Bierman, E. L. 1972. *Metabolism* 21:1009–19
53. Havel, R. J., Kane, J. P. 1973. *Proc. Nat. Acad. Sci. USA* 70:2015–19

Copyright 1975. All rights reserved

DISORDERS OF THE LOWER ESOPHAGEAL SPHINCTER

❖7159

Robert S. Fisher, M.D.
Gastrointestinal Section, Department of Medicine, Temple University Health Sciences Center, Philadelphia, Pennsylvania 19140

Sidney Cohen, M.D.
Gastrointestinal Section, Department of Medicine, University of Pennsylvania at the Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania 19104

INTRODUCTION

The junction of the esophagus and stomach in man is marked by a zone of elevated pressure which has been called the lower esophageal sphincter (LES) (1, 2). The purpose of this sphincter mechanism is to prevent regurgitation of gastric contents from a positive pressure cavity into the esophagus which maintains a pressure negative to the atmosphere (3, 4). This sphincter must not only prevent gastroesophageal reflux, but also it must open or relax when called upon to allow passage of food into the stomach (5).

A number of diseases in which the LES fails as an antireflux mechanism or fails to relax properly during swallowing have been recognized. Clarification of the pathophysiology of these diseases and an approach to newer forms of therapy depend upon understanding the basic physiology of LES function. The purpose of this review is threefold: 1. to discuss recent concepts involving the regulation of LES strength and LES relaxation; 2. to define the basic abnormalities recognized in disorders of LES function; and 3. to explore the physiological basis of new approaches to pharmacological and surgical therapy.

THE CONCEPT OF A PHYSIOLOGICAL SPHINCTER MECHANISM

The inability of anatomists and pathologists to convincingly demonstrate an anatomic sphincteric structure in the distal esophagus (6, 7) has produced skepticism and misunderstanding concerning the role of the LES in the pathophysiology of esophageal diseases. The concept of a physiological sphincter is based upon recognition that the distal portion of the esophagus, although anatomically indistinct, functions differently than the proximal body of the esophagus. In contrast to the LES which is tonically contracted in the basal state but relaxes with swallowing or esophageal distention, the esophageal body is at rest basally but contracts in response to swallowing or esophageal distention. This reciprocal activity of the distal esophagus compared to the body of the esophagus is the major evidence for a physiological sphincter mechanism. Abnormalities of the complex myogenic and neural mechanisms that determine the normal behavior of this sphincter are responsible for various esophageal disorders.

Several specific properties of the distal esophagus provide evidence for the existence of a "physiologic" LES. First, the circular smooth muscle of the LES has different mechanical properties when studied in vitro (8, 9). LES muscle, as compared to esophageal muscle, requires less passive stretch to achieve its length at which maximal active tension is developed in response to a stimulus. Thus, the sphincter muscle achieves greater tension at a shorter muscle length than the adjacent esophagus (10). This myogenic property may be an important factor in determining sphincter function. Second, LES muscle, in vitro, contracts at a lower threshold dose and achieves a greater maximum active tension when stimulated with the hormone, gastrin, and the neural transmitters, acetylcholine and norepinephrine (10, 11). Third, LES muscle shows a specific neurogenic inhibitory response not seen in the body of the esophagus (9, 12). In summary, the following factors determine the physiological characteristics of LES function. LES closure is a property of muscle which contracts with minimal stretch and develops greater force with less neurogenic or hormonal stimulation than adjacent esophageal and gastric muscle. Sphincter relaxation is a function of a localized inhibitory innervation. Thus, sphincter closure and sphincter relaxation are the major physiological properties which characterize the muscle and innervation of the distal esophagus.

PHYSIOLOGY OF LES FUNCTION

The regulation of LES pressure can be divided into three headings: 1. basal LES pressure; 2. adaptive alterations in LES pressure; and 3. LES relaxation (Table 1).

Basal LES Pressure

In the basal state in man, the LES is a 2–5 cm zone of elevated resting pressure. In normals this resting pressure when measured using infused, open-tipped, recording catheters is 12–30 mm Hg above the intra-abdominal pressure level (1). LES pressure, as measured with infused catheters, is an excellent index of the strength of the gastroesophageal antireflux mechanism (4). Because different methods of recording pressure from the LES give variable results, it is important that the method be chosen carefully (13).

The strength of the gastroesophageal antireflux mechanism resides in the intrinsic LES. Simple LES displacement, as seen in sliding hiatal hernia, does not diminish

Table 1 Regulation of LES pressure

- Basal LES Pressure
  - LES muscle properties
  - Gastrin
  - Alpha-adrenergic stimulation
  - Vagal tone
- Adaptive Alterations In LES Pressure
  - Increase
    - Gastrin
    - Cholinergic reflexes
  - Decrease
    - Secretin
    - Cholecystokinin
    - Glucagon
    - Prostaglandins
    - Decreased gastrin
- LES Relaxation
  - Purinergic nerves
  - Prostaglandins

LES strength (14). Our understanding of the factors regulating basal LES strength has been derived from studies designed to selectively remove possible contributing physiological factors. The majority of these studies have been performed in the opossum, chosen because its smooth muscle esophageal body and LES are similar to those in man. Basically, three major observations have been made. First, selective inactivation of the hormone gastrin by antigastrin antiserum gives a 70 to 80% reduction in basal LES pressure (15). Also, in man nearly a 70% decrease in LES pressure has been observed when endogenous gastrin release was inhibited by gastric acidification (16). However, significant reduction in LES pressure during gastric acidification has not been observed by some investigators (17). Second, specific adrenergic nerve destruction with 6-hydroxydopamine or the administration of the alpha-adrenergic antagonist, phentolamine, produces a modest LES pressure reduction of about 20 to 30% (18). Third, although bilateral cervical vagal section does not alter LES tone in the opossum (8), there is a suggestion that vagal tone may be of some importance in other animals (19, 20). At present, it seems that the genesis of basal LES tone is due to a combination of sphincter muscle mechanical properties, gastrin, alpha-adrenergic stimulation, and vagal tone.

Adaptive Alterations in LES Pressure

Basal LES pressure has been shown to change with various physiological stimuli. These changes, as presently understood, can be divided into two major categories: 1. hormonal response to food ingestion; and 2. reflex changes in response to alterations in intra-abdominal pressure.

It has now been demonstrated that the human LES responds to each of the gastrointestinal hormones, gastrin (10, 21), secretin (22, 23), cholecystokinin (24, 25), and glucagon (26). Only gastrin increases LES pressure whereas each of the other hormones reduces it. To date, most observations on the effects of the gastrointestinal hormones on LES pressure are based upon observations made during their exogenous administration. However, the physiological effect of several hormones has been confirmed in man. Endogenous gastrin released by a protein meal, beef hydrolysate, or glycine has been shown to augment LES strength (27). Endogenous cholecystokinin released by fat decreases LES pressure (28, 29). The effect of secretin is more complex in that its endogenous release inhibits the effect of endogenous or exogenous gastrin without altering the basal LES tone (22). These studies would suggest that hormonal factors play a role in regulating LES pressure during feeding. Additionally, hormonal factors may also be operative in the production of esophageal symptoms related to the ingestion of specific types of food (28).

Another important LES adaptive response is that seen with changes in intra-abdominal pressure. In normals, it has been shown that an increase in intra-abdominal pressure is associated with an even greater increase in LES pressure (14, 30). This important response is probably operative in the prevention of gastroesophageal reflux during exercise, bending, or a Valsalva maneuver. This reflex response is believed to be mediated by afferent fibers carried in the vagal trunk and can be reduced by vagotomy (19) or the administration of atropine (20). Also, this reflex has been found to be abnormal in patients with an incompetent LES (14). In these patients, the LES pressure increase is exceeded by the change in intra-abdominal pressure, a situation promoting reflux of gastric contents into the esophagus.

LES Relaxation

The LES relaxes in response to swallowing or proximal esophageal distention. Relaxation is an active neurogenic inhibitory process that is mediated by specialized nonadrenergic, noncholinergic nerves with specific localization in the distal esophagus (9, 12). The nonadrenergic inhibitory innervation has been classified recently as a "purinergic" system working through the neural transmitter, adenosine triphosphate, or a related purine nucleotide (31). Although our understanding of the purinergic innervation of the gastrointestinal tract is quite sparse, further developments in this area are of great importance. Incomplete or partial LES relaxation, as mediated by the purinergic system, is a major LES defect in patients with achalasia (32) and in some patients with symptomatic diffuse esophageal spasm (33). As will be discussed later drugs that may augment purinergic nerve-mediated inhibition may increase the magnitude of LES relaxation.

Further evidence for a specific nucleotide-mediated innervation for LES relaxation has recently been demonstrated using prostaglandin E_1 (34). It has been shown that exogenously administered prostaglandin E_1 relaxes the LES through a second messenger system working through cyclic AMP. These studies raise the possibility that not only a nucleotide but also prostaglandin may be involved in LES relaxation. A twentyfold elevation of prostaglandin E concentration has been reported in the frog esophagus in response to electrical stimulation (35).

Since LES relaxation is a function of purinergic nerves, many widely used pharmacological agents, fortunately, do not diminish its magnitude. Anticholinergics, alpha- and beta-adrenergic antagonists, and ganglionic blocking agents do not alter LES relaxation (34) and, therefore, do not disrupt esophageal emptying.

TESTS OF LES FUNCTION

Heartburn, waterbrash, regurgitation, and bilious vomiting are all manifestations of gastroesophageal reflux resulting from an incompetent LES. In contrast, a hypertensive LES which relaxes incompletely may lead to esophageal distention and cause symptoms of dysphagia and odynophagia. Because LES dysfunction can be responsible for such a variety of symptoms, many tests have been devised to evaluate its function. These can be divided into two major categories: indirect and direct. Although there is general agreement that direct testing is most informative about the LES itself, there is much controversy over which test correlates best with symptoms.

The indirect techniques for evaluating LES function are barium esophagography, barium cine-esophagography, endoscopy, biopsy, acid-perfusion, acid-reflux, and acid-clearance testing. Most commonly employed is the conventional barium swallow. After ingestion of a barium bolus, its progression through the esophagus into the stomach is followed fluoroscopically and appropriate spot films are taken before and after the gastroesophageal pressure gradient is manipulated by various maneuvers. The advantages of barium esophagography are that it is a routine procedure that is easily performed, and necessary equipment is readily available. The disadvantages are total dependence upon the fluoroscopist for interpretation and choice of spot films, and insensitivity to subtle abnormalities. Gastroesophageal reflux is detected in only 10–25% of symptomatic patients by this technique (36).

Employing image intensification and televised fluoroscopy, the fluoroscopic image can be recorded on motion picture or videotape. Cine-esophagography is preferable to conventional studies in that a permanent record of the entire examination is available for review, and greater diagnostic accuracy can be achieved by using slow motion or frame-by-frame analysis. Although reflux can be detected in 40–67% of patients (37), the test is quite expensive. Some investigators taking note of the abnormal esophageal contractions which often accompany acid reflux have employed acid barium for these studies (38). The acid barium swallow seems to be a useful screening test for reflux, but the incidence of false positives is unacceptable. Another test employing radiographic techniques is the "water-siphon" test (39). Here gastroesophageal reflux is looked for during repetitive swallowing. Since normally the LES relaxes and remains relaxed with repetitive swallowing, this test also has an unacceptable incidence of false positives (36).

An erythematous, eroded, or ulcerated mucosa as detected by endoscopic examination is diagnostic of esophagitis (40). In contrast, the esophageal mucosa may appear normal to inspection with underlying inflammation only detectable by histologic examination of biopsy specimens. Tissue examination is the most accurate means of detecting esophagitis. Preferably, specimens should be obtained with suction biopsy instruments which allow not only adequate sampling of the surface epithelium and lamina propria but also proper orientation of the specimen (41). Although not ideal, biopsy specimens may be obtained during endoscopy (42). Endoscopy, biopsy, and cytological examination of tissue are the only sure methods of excluding malignancy.

Since the symptoms of LES incompetence are generally attributed to reflux of acid gastric material into a sensitive esophagus, several procedures employing instillation of 0.1N HCl have been devised. The acid-perfusion test of Bernstein has achieved the widest acclaim (43). Here, classical reflux symptoms are produced when 0.1N HCl is dripped into the esophagus. Symptoms disappear when HCl is replaced by normal saline. Many observers have reported that a positive Bernstein test correlates best with symptoms (44). The acid-reflux test (45) employs a pH probe implaced in the distal esophagus, 2–5 cm proximal to the LES. After acid is instilled into the stomach, intra-abdominal pressure is increased. A drop in distal esophageal pH below 4 denotes a positive test. Another variant is the acid-clearing test (46). After the stomach is filled with acid, a pH probe is slowly withdrawn from the stomach, across the LES, to a position about 2–5 cm above the LES. The number of swallows necessary to clear acid from the distal esophagus (restore $pH > 4$) is recorded. Normally, 4–10 swallows are necessary to clear acid from the distal esophagus. As an alternative, acid may be instilled directly into the distal esophagus with the pH probe in place 2–5 cm above the LES. These acid tests are positive in 75–95% of symptomatic patients (44). The abundance of diagnostic tests to detect gastroesophageal reflux attests to the absence of a convenient, accurate procedure. Clearly, what is needed is an easy, accurate method of quantitating reflux and relating it to both gastroesophageal pressure gradient and LES pressure.

Intraluminal manometry constitutes the only direct method for evaluating LES function. Code and Ingelfinger initially employed balloon manometry to measure LES pressure and relaxation as a recording device was slowly withdrawn from the stomach into the esophagus in 1-cm increment (1, 47). Studies using open-tipped, water- or air-filled tubes in place of balloons, with either end-or-side orifices, followed. Presently, open-tipped, infused tubes with side orifices are considered to give the best measurements of LES pressure (3). Recently, direct-recording microtransducers have been employed also (48).

DISORDERS OF LES FUNCTION

Disorders of the LES can be grouped under two headings: decreased LES pressure and increased LES pressure with or without normal LES relaxation (Table 2).

Decreased LES Pressure

LES INCOMPETENCE WITH OR WITHOUT HIATAL HERNIA For many years, it has been known that heartburn and other reflux symptoms occur frequently in the presence of a sliding hiatal hernia. Only recently, however, heartburn and unequivocal reflux symptoms have been described in the absence of any other clinical disorder. This condition is called primary LES incompetence (49, 50). Reflux in patients

Table 2 Disorders of the lower esophageal sphincter

Decreased LES Pressure
1. LES Incompetence with or without hiatal hernia
2. Progressive systemic sclerosis (Scleroderma)
3. Pregnancy
4. Post-vagotomy
5. Neuropathy (diabetes or alcoholic)
6. Infantile chalasia

Increased LES Pressure With or Without Normal LES Relaxation
1. Achalasia
2. Chagas' disease
3. Diffuse esophageal spasm
4. Malignancy

with hiatal hernia has been ascribed to various anatomical abnormalities (51): loss of the normal cardioesophageal angle of His, widening of the diaphragmatic hiatus, loss of the normal gastric rosette formed by interdigitating gastric mucosal folds, disruption of the phrenoesophageal ligament, conversion of the distal paraesophageal pressure from intra-abdominal pressure to intrathoracic pressure, and dysfunction of the LES. Much clinical and experimental evidence, accumulated over the past two decades in situations where accessory resistance factors have been disrupted, suggests that the LES is the major determinant of gastroesophageal competence (14). In normal individuals, the LES can prevent reflux despite intragastric pressures in excess of 10–15 mm Hg. Normally LES pressure ranges between 12 and 30 mm Hg above intragastric pressure. Gastroesophageal reflux occurs when LES pressure drops below 10 mm Hg. Nevertheless, reflux has occasionally been described in patients with normal LES pressures. There are several possible explanations. First, intra-abdominal pressure may become so high as to overwhelm the esophago-gastric junction resistance and produce reflux. This is unlikely since it has been shown that increments in intragastric pressure are normally accompanied by even greater increases in LES pressure (30). Second, reflux may be an intermittent phenomenon corresponding to intermittent decreases in LES pressure. Currently, our diagnostic procedures do not employ simultaneous reflux and manometric measurements.

Since the majority of patients with hiatal hernia have no symptoms which can be ascribed to the hiatal hernia, an explanation for reflux other than the disturbed anatomy has been sought. Although there is agreement that, in LES incompetence with or without hiatal hernia, LES dysfunction is the major cause of reflux, the pathophysiology of LES dysfunction remains an enigma. Recently, however, evidence has been presented to suggest that both decreased LES stimulation and decreased LES responsiveness are present. An impairment in the antral gastrin-release mechanism, as evidenced by decreased serum gastrin concentrations (52) and decreased integrated gastrin response (53) following stimulation of gastrin release, has been demonstrated in patients with LES incompetence. In addition, direct LES stimulation by synthetic pentagastrin or by cholinergic agents has produced decreased absolute LES responses in these same patients (54). Furthermore, both basal LES pressure and LES responsiveness to hormonal and cholinergic stimulation seem to return to normal following surgical repair (55–57). This holds for both primary LES incompetence and hiatal hernia. These observations suggest that the condition may be reversible. Although a definitive explanation for this last phenomenon is wanting, we suggest that a change in the length-tension characteristics of the LES circular muscle is of importance. The active tension produced by a muscle in response to a fixed stimulus is in part dependent upon the length of the muscle at the time of stimulation. Each muscle has its own length of optimal tension development (L_0). The striated muscles in the body with their fixed origins and insertions are set at their L_0s. The same probably is true of both cardiac and smooth muscle. Perhaps, in LES incompetence, there is an alteration in the length of LES circular muscle away from its L_0 and, thereby, a change in both basal LES pressure and responsiveness to stimulation. Surgical repair might inadvertently restore the LES to L_0 and improve its function. Certainly, further studies in this area are needed.

The objectives of medical treatment of LES incompetence can be divided into several categories: decrease gastroesophageal pressure gradient; decrease gastric hydrochloric acid secretion; and protect the esophagus from refluxed HCl. The gastroesophageal pressure gradient can be minimized by having the patient avoid the recumbent position by the use of pillows, blocks under the legs at the head of the bed, or even by the use of a hospital bed. In addition, he should be instructed to avoid maneuvers, such as lifting heavy objects, which increase intra-abdominal pressure. Marked weight reduction undoubtedly helps some obese patients, perhaps by removing the fat panniculus from pressing on the stomach. Hydrochloric acid can be minimized by avoiding acid-secretory stimulants such as caffeine or alcohol and eating a bland diet. Anticholinergic agents, although they may decrease acid secretion, may also decrease LES pressure (20). Various agents are employed to protect the esophagus from refluxed acid. The mainstay is antacid therapy which should neutralize the acid before it reaches the esophageal mucosa. Another indirect benefit derived from antacids may be increased LES pressure (58). This possibly is due to antral gastrin release or cholinergic reflexes. However, there is some doubt whether antacid therapy changes LES pressure at all (17, 59). Patients should be instructed to chew their food well since this might increase salivary volume and bicarbonate (60). Lastly, some agents such as alginic acid may create a mechanical barrier to decrease gastroesophageal reflux (61). Some newer methods of treating LES incompetence are on the horizon and will be discussed later.

The indications for surgical treatment of LES incompetence are as follows: intractability to conservative medical and pharmacologic therapy, severe hemorrhage produced by reflux esophagitis, and esophageal stricture formation producing partial or total obstruction. Excellent results have recently been reported for the surgical procedures popularized by Belsey (62), Nissen (63), and Hill (64).

PROGRESSIVE SYSTEMIC SCLEROSIS (SCLERODERMA) Visceral involvement in systemic sclerosis is common and may occur in the absence of typical skin involvement (65). Usually, esophageal involvement is in association with Raynaud's phenomenon (66). From one third to one half of all patients with systemic sclerosis complain of heartburn or dysphagia at some time during the course of the disease (67). At times these symptoms may be most distressing. Radiographically, impaired motility or aperistalsis of the distal two thirds of the esophagus has been well described. Also, copious gastroesophageal reflux of barium is frequently demonstrated across a widened, patulous gastroesophageal junction (68). Intraluminal manometry is even more sensitive in detecting these abnormalities. The distal two thirds or smooth muscle portion of the esophagus is characterized by feeble, often nonprogressive contractions. Also, a decreased or absent LES high pressure zone may be present (69). The upper one third of the esophagus and upper esophageal sphincter are usually normal. Because of copious gastroesophageal reflux and a defective esophageal clearance mechanism, severe esophagitis and esophageal strictures are not unusual.

Recently, the visceral smooth muscle involvement of systemic sclerosis has been ascribed to a neuromuscular defect and not to the presence of sclerosing, proliferating collagen (70). At postmortem examination, the smooth muscle is atrophic but collagen deposition and fibrosis are not prominent (71). Other observations have also suggested a neuromuscular basis for this disorder. Functional derangements may precede pathological changes as noted by comparison of manometric esophageal recordings and postmortem examination for muscle atrophy by light microscopy (71). Esophageal dysfunction correlates closely with Raynaud's phenomenon, suggesting that a common neurogenic abnormality may cause both disorders (66). Intra-arterial reserpine improves peripheral capillary blood flow in patients with Raynaud's phenomenon and has reversed some of the esophageal abnormalities associated with systemic sclerosis (72), suggesting that in early disease, muscle function is preserved. Lastly, in patients with early systemic sclerosis, the LES response to direct muscle stimulation by methacholine was intact while the response to gastrin and edrophonium, an acetylcholinesterase inhibitor, was diminished (70). These observations suggest an underlying neurogenic mechanism. Further word on this subject is expected.

PREGNANCY Heartburn, frequently distressing, occurs in one third of pregnant women (73). Presumably, it is related to reflux of acidic gastric contents across the LES causing esophagitis. Usually, these symptoms resolve spontaneously after pregnancy has terminated. In some women, however, stricture formation and esophageal obstruction have been reported. Formerly, heartburn of pregnancy had been attributed to elevated intragastric (intra-abdominal) pressure. However, it has now been demonstrated that, normally, an increase of intragastric pressure is accompanied by an even greater increment of LES pressure, thereby preserving LES competence (30). Recently, diminished resting LES strength has been reported in pregnant women during the third trimester. This returns to normal after delivery (74, 75). The disorder might be hormonal, neural, muscular, or a combination of these. Progesterone is known to be a smooth muscle relaxant (76), but its effect on the LES has not been measured. The observations that gastrin is a mediator of LES tone and that cholinergic stimulation can increase LES pressure suggest that a further look at the pathophysiology of heartburn of pregnancy is indicated.

POST-VAGOTOMY Heartburn and regurgitation of bilious material is not an infrequent complaint following the performance of truncal vagotomy. A postoperative decrease in LES pressure response has been recorded (10). A role for cholinergic nerves has been proposed both for maintaining basal LES pressure and for increasing LES response to increased intragastric pressure. Medical vagotomy, i.e. administration of atropine sulfate, both decreases basal LES tone and depresses LES responsiveness in man (20). These data alone do not firmly establish a role for cholinergic fibers since gastrin has been shown to work at the LES by releasing acetylcholine from postganglionic cholinergic sites (21).

ALCOHOLIC AND DIABETIC NEUROPATHY Both of these disorders affect the autonomic nervous system, are poorly understood, and are frequently associated with nonspecific esophageal motor disorders (77, 78). Occasionally, a claim of LES incompetence has been made. However, newer manometric techniques have not been employed to study the LES in either alcoholic or diabetic neuropathy.

CHALASIA OF INFANCY That infants regurgitate easily has been known for some time. Until recently, the explanation for this observation was not clear. It has now been shown that infants have reduced LES pressures as compared to adults (79). In recent studies on the opossum it was also found that LES pressure was markedly reduced in infants and gradually reached normal levels as the animals reached adulthood (80). The reduction in LES pressure was attributed to two causes, decreased muscle mass and a specific, diminished LES muscle responsiveness to gastrin. As the animals matured, LES pressure increased and LES responsiveness to gastrin improved. These observations indicate that infantile chalasia is a normal developmental state and that antireflux surgery in infants should be delayed until one is certain that sphincteric function has matured. The relationship of infantile chalasia to LES incompetence in adults is not known.

Increased LES Pressure With or Without Normal LES Relaxation

ACHALASIA This disease has been called by many names including aperistalsis, megaesophagus, and cardiospasm. Actually, achalasia too is a misnomer since technically achalasia means a total lack of relaxation of the LES after swallowing. We now know that in achalasia the LES can usually relax, but only partially (32). Classically, achalasia is characterized by several manometric abnormalities. First, there is absence of purposeful, aborally progressive, lumen-obliterating contractions throughout the entire esophagus (81, 82). Consequently, in the absence of gravity or a hydrostatic pressure head, a food bolus is not propelled over the length of the esophagus. Second, the basal pressure of the LES itself is usually increased above normal (83). Third, the hypertensive LES does not relax normally following a swallow or proximal esophageal distention (32). In normal subjects, the LES relaxes 95–100% towards the intragastric pressure level following appropriate stimulation. Therefore, there is little or no resistance to aboral movement of a food bolus into the stomach. However, in achalasia LES relaxation is absent or incomplete, leaving a significant residual pressure which acts as a barrier to orderly emptying of the esophagus. In effect, there is a distal mechanical obstruction of the esophagus produced by a tonically contracted LES which fails to relax completely (84) in addition to esophageal pump failure due to absence of orderly peristalsis.

The symptoms of achalasia are directly related to esophageal obstruction. The most common complaint, occurring in almost all patients, is dysphagia for both solids and liquids (85). In addition, these patients frequently complain of intermittent retrosternal or interscapular pain, probably related to esophageal distention or tertiary esophageal contractions. Failure of food particles to leave the esophagus often produces nocturnal aspiration or possibly regurgitation. Weight loss is usually mild and may occur because of fear of eating. True anorexia is rare. Odynophagia can occur in achalasia but is more typical of diffuse esophageal spasm. True heartburn virtually excludes this diagnosis.

Diagnosis is made by X ray and/or manometric techniques. In late achalasia barium esophagography typically shows a dilated, flaccid, bag-like esophagus with a smoothly tapered, symmetrical, beak-like distal segment which represents the tonically closed LES (86). In some cases the dilated esophagus can actually be seen on chest X ray. Early esophageal manometric testing may detect disease before esophageal dilatation occurs. Cine-barium-esophagography may show absence of esophageal peristalsis with a typical distal segment in the absence of a dilated esophagus. Esophagogastroscopy should be performed in all patients suspected of having achalasia. There seems to be an increased incidence of esophageal carcinoma, usually involving the mid-esophagus, in long standing achalasia (87). Also, tumors of the cardia can simulate achalasia on X ray (88, 89). In the majority of patients with achalasia the endoscope is easily passed through the hypertensive LES into the stomach allowing thorough evaluation of the cardioesophageal junction. The "Mecholyl test" has also been used as a diagnostic test for achalasia (90, 91). Following the subcutaneous administration of methacholine (Mecholyl®) the esophagus contracts violently. The production of severe crushing substernal or back pain, the orad propulsion of barium by the esophagus, and an increase of baseline intraesophageal pressure $\geq$ 25 mm Hg constitute a positive test (92). Unfortunately, a positive Mecholyl test is not specific for achalasia. Positive tests have been recorded in diffuse esophageal spasm (93), Chagas' disease (94), and cancer of the cardia (99).

Most agree that achalasia is a primary neuromuscular disease. An exaggerated esophageal response to cholinergic stimulation was observed years ago and forms the basis for the Mecholyl test. Also, an exaggerated LES response to gastrin administration has recently been recorded (95). These observations are compatible with Cannon's law of denervation supersensitivity (96). Further support of a neurogenic basis for achalasia is provided by the following observations. Although atrophy and degeneration of esophageal smooth muscle are usually present, muscle changes are minimal in some cases. In the majority of patients with achalasia, histopathologic study shows an absence or a decrease of parasympathetic ganglion cells within the myenteric plexus of the esophagus (97). However, Cassella et al reported 9 out of 28 patients with normal myenteric plexuses and suggested a relation to duration of disease (98). Electron microscopic examination of vagal nerve fibers in patients with achalasia often revealed slight changes suggestive of Wallerian-type degeneration. Further examination of the dorsal motor nucleus of the vagus in two patients with achalasia who had died showed a markedly reduced number of dorsal motor nuclear ganglion cells. Therefore, it has been suggested that achalasia begins as a disease of the extraesophageal nervous system, progressing to deplete the esophageal myenteric plexus of ganglia, and to produce a denervation supersensitivity of the esophagus and LES.

A precise etiologic agent producing achalasia has not been found. Because of this, treatment remains symptomatic and is aimed at relieving distal esophageal obstruction. The choice between surgical myotomy (99) (Heller procedure) and forceful pneumatic or hydrostatic dilatation (100) depends largely upon local talent. We prefer the nonsurgical approach for several reasons. First, symptoms correlate closely with LES basal pressure (84), and we are usually able to decrease LES pressure to normal levels with dilatation. Second, esophageal perforation is unusual (101). Third, surgery can always be performed if other therapeutic measures fail. Fourth, significant gastroesophageal reflux occurs in 20–40% of patients following a Heller myotomy (102), and resultant stricture formation is not unusual. A pharmacologic approach to decreasing LES pressure is currently being actively investigated.

CHAGAS' DISEASE Chagas' disease, the usual name given to trypanosomiasis americana, is widely distributed in Latin America, where it has a high mortality in children and produces much incapacity among adults. Cases have not infrequently been described in the southern United States. Its parasitic agent, *Trypanosoma cruzi* can damage the heart, esophagus, small intestine, colon, and sometimes the bronchi, ureter, and urinary bladder. The anatomical and functional disorders of Chagas' disease are mainly the result of damage to the parasympathetic ganglia of the autonomic nervous system (103, 104). Neurolysis follows rupture of parasitic pseudocysts with the liberation and disintegration of leishmanial forms of the parasite close to the neurones. Chagas' disease involving the esophagus is difficult to differentiate from achalasia. The clinical symptoms and radiographic appearance are quite similar. One would predict similar manometric findings as well (94). Chagas' disease of the esophagus is suspected when the patient is from an endemic area or when other portions of the gastrointestinal tract are similarly involved. Final diagnosis can be made by specific serologic testing. Here, as in achalasia, there is no specific treatment. Again, efforts are aimed at relieving lower esophageal obstruction.

DIFFUSE ESOPHAGEAL SPASM Diffuse esophageal spasm is a motor disorder characterized by painful swallowing, and impairment of the esophageal transport process (105, 106). These patients complain of odynophagia, retrosternal pain, and dysphagia for both liquids and solids. Radiographic features are tertiary contractions which range from low to very high amplitude and have prompted the eponyms of curling, rosary-bead, and corkscrew esophagus (107–109). Esophageal manometry characteristically shows high amplitude, repetitive, simultaneous (spastic), nonperistaltic contractions in the esophageal body. Some normal peristaltic activity is always present. LES function has been considered normal in patients with diffuse esophageal spasm. Recently, however, it has been shown that up to 33% of these patients have abnormal LES function (33). The LES may have impaired relaxation with or without an increased LES pressure. This subgroup has been previously called either "vigorous" achalasia or diffuse esophageal spasm with hypertensive LES. The therapeutic approach to diffuse esophageal spasm is usually aimed at the hypermotile esophagus. Anticholinergic agents have achieved some success. Topical anesthetics have, in general, been unsuccessful. When the patient complains of severe dysphagia, and LES dysfunction is prominent, procedures or medications intended to decrease LES pressure may also be helpful.

The cause of diffuse esophageal spasm is unknown. There are some who believe that achalasia and diffuse esophageal spasm are different ends of the same disease spectrum (110). However, in diffuse esophageal spasm electron microscopic studies have shown a picture quite distinct from achalasia (111). Here only minimal focal morphologic changes are usually seen in esophageal smooth muscle cells and the myenteric plexus, accompanied by generalized degenerative changes in the afferent vagal fibers. In contrast, in achalasia profound changes are usually present in esophageal smooth muscle cells with only minimal focal changes in vagal nerve fibers. There is some evidence to suggest that diffuse esophageal spasm represents a disorder of esophageal "purinergic" nerves (33).

MALIGNANCY There are several case reports in the literature where malignancy involving the gastric cardiac has been mistaken for achalasia. Clinical symptoms, X rays, and manometric studies can all suggest a diagnosis of achalasia (88). Even positive Mecholyl tests have been reported (89). These disorders are thought to be caused by malignant infiltration of the myenteric plexus. If endoscopy is performed on all of these patients, this diagnosis should not be missed.

PHARMACOLOGY OF THE LES

Now that we understand the physiologic mechanisms regulating the LES, a more rational approach to treatment of its dysfunction is possible. The pharmacologic approach to treating disorders of the LES signals the beginning of a new era. This section is an attempt to outline the present status of this highly investigational area. We have chosen to divide this section into discussions of agents which increase LES pressure, decrease LES pressure, and augment LES relaxation (Table 3).

Increase LES Pressure

ACETYLCHOLINE AND BETHANECHOL Atropine has been demonstrated to decrease basal LES pressure and decrease the LES response to both increased intragastric pressure (20) and gastrin stimulation (56). In 1960, Bettarello et al noted increased LES pressure in both normal subjects and patients with gastroesophageal reflux after injection of a cholinergic agent (112). Employing newer manometric

Table 3 Pharmacology of LES

Agent	Mechanism of Action
Increase LES Pressure	
Acetycholine Bethanechol	Muscarinic effect on ganglia and muscle
Eserin Neostigmine Edrophonium	Anticholinesterase action
Phenylephrine Norepinephrine	Alpha-adrenergic stimulation of muscles
Metoclopramide	Direct muscle effect; acetylcholine release
Antacids	Antral gastrin release (?); cholinergic reflex
Decrease LES Pressure	
Atropine Propantheline Dicyclomine	Anticholinergic action
Isoproterenol	Beta-1 adrenergic stimulation
Carbuterol Salbutamol	Beta-2 adrenergic stimulation
Prostaglandin E_1, E_2, A_2	Increases cyclic AMP
Theophylline	Phosphodiesterase inhibition
Lipids (oral)	Release cholecystokinin from small bowel mucosa
Augment Relaxation	
Dipyridamole	Adenosine uptake inhibition Phosphodiesterase inhibition
Lidoflazine	Adenosine uptake inhibition

techniques, these observations have been confirmed and extended for oral bethanechol (113). Recently, a controlled, double-blind therapeutic evaluation of orally administered bethanechol has suggested its efficacy in treating reflux patients (114). Presumably, the observed activity of cholinergic stimulation is due primarily to muscarinic effects on muscle receptors and, perhaps, parasympathetic ganglia.

ESERINE, NEOSTIGMINE, AND EDROPHONIUM Since acetylcholine increases LES pressure, agents which increase its concentration at ganglia or muscle receptors will likewise increase pressure (70). Eserine, neostigmine, and edrophonium inhibit acetylcholinesterase and prevent local acetylcholine metabolism. For them to be effective, acetylcholine must be available at the receptor sites.

PHENYLEPHRINE AND NOREPINEPHRINE That alpha-adrenergic stimulation caused LES circular muscle to contract, in vitro, has been known since 1968 (115).

However, only recently, studies performed on the opossum, in vivo, have shown that LES pressure can be increased about 64% and 42% by intravenous administration of norepinephrine and phenylephrine, respectively (18). Unfortunately, blood pressure and pulse rate also increase dramatically with these agents. A role for alpha-adrenergic stimulation in maintaining basal LES pressure has been demonstrated in the opossum. No studies have yet been reported on human subjects.

METOCLOPRAMIDE This agent has been shown to have a number of actions on gastrointestinal motility (116). Metoclopramide clearly raises the pressure of the LES above basal levels when administered either intravenously (117) or orally (118). This increased LES pressure has been related to a decreased frequency of reflux episodes. The effect of metoclopramide is partially blocked by atropine. It seems to act directly on the muscle and by releasing acetylcholine.

ANTACIDS As mentioned previously, antacids undoubtedly are useful to prevent and treat symptoms of gastroesophageal reflux. Certainly, they buffer acid within both the stomach and the esophagus. In the past, increased LES pressure after gastric deacidification had been attributed to antral gastrin release (57). Recently, the latter mechanism has been challenged (58, 59).

Decrease LES Pressure

ATROPINE, PROPANTHELINE, AND DICYCLOMINE These agents all act by blocking cholinergic receptor sites. Although there is no doubt that they can decrease LES pressure if given in high enough dosage, unacceptable side effects may be produced at these levels.

ISOPROTERENOL, CARBUTEROL, AND SALBUTAMOL Recently, it has been shown that stimulation by beta-adrenergic agonists decreases LES pressure (119). These agents may act by stimulating adenyl cyclase. Unfortunately, isoproterenol, in addition to decreasing LES pressure by 50–60%, also greatly decreased the blood pressure and increased the heart rate. Specific beta-2 adrenergic stimulation with Carbuterol and Salbutamol also significantly decreased LES pressure, but with less effect on blood pressure and heart rate. These agents might be useful in treating the hypertensive LES in achalasia.

PROSTAGLANDINS E_1, E_2, and A_2 Prostaglandins have now been demonstrated to significantly decrease resting LES pressure not only in the opossum but also in normal subjects and in patients with achalasia (34).

PAPAVERINE AND THEOPHYLLINE This is a third category of pharmacologic agent which decreases LES pressure by increasing cyclic AMP (34). In this case phosphodiesterase is inhibited preventing the metabolism of cyclic AMP. Neither agent has been evaluated therapeutically.

LIPIDS (ORAL) Lipids decrease LES pressure (28, 29). This phenomenon is believed to be mediated by cholecystokinin release from the small intestinal mucosa. Cholecystokinin may act by competing with gastrin at the LES receptor (25).

Augment LES Relaxation

One approach to the treatment of a hypertensive LES which fails to relax would be to decrease resting pressure. Pneumatic dilatation, the Heller myotomy, and the pharmacologic agents described above all decrease resting LES pressure. Another approach would be to improve LES relaxation without affecting resting LES pressure. If this were accomplished, the residual pressure barrier to esophageal emptying would be removed or decreased.

DIPYRIDAMOLE AND LIDOFLAZINE These agents block the intracellular uptake of adenosine, thus potentiating purinergic nerve function. Both dipyridamole and lidoflazine have recently been shown to potentiate the maximal inhibitory responses of LES muscle to electrical stimulation during muscle contraction produced by either acetylcholine or KCl depolarization (120). Dipyridamole was more potent, possibly due to its additional phosphodiesterase-inhibitory properties. These drugs have not been tested in man.

ACKNOWLEDGMENTS

This work was supported by Research Grants 5 RO1 AM 16280-02 and RR-054-17 from the National Institutes of Health and 1 RO1 AA00656-01 from the National Institutes on Alcohol Abuse and Alcoholism. Dr. Cohen is supported by a Research Career Development Award 5 KO4 AM 70 576-01 from the National Institutes of Health.

Literature Cited

1. Fyke, F. E., Code, C. F., Schlegel, J. F. 1956. *Gastroenterologia* 86:135–50
2. Botha, G. S. M., Astley, R., Carre, J. J. 1957. *Lancet* 272:659–61
3. Winans, C. S., Harris, L. D. 1967. *Gastroenterology* 52:773–78
4. Cohen, S., Harris, L. D. 1970. *Gastroenterology* 58:157–62
5. Fleshler, B. et al 1958. *J. Appl. Physiol.* 12:339–42
6. Allison, P. R. 1951. *Surg. Gynecol. Obstet.* 92:419–51
7. Higgs, B., Kerr, F. W. L., Ellis, F. H. 1965. *J. Surg. Res.* 5:503–7
8. Christensen, J., Freeman, B. W., Miller, J. K. 1973. *Gastroenterology* 64:1119–25
9. Christensen, J., Conklin, J. L., Freeman, B. W. 1973. *Am. J. Physiol.* 225:1265–70
10. Lipshutz, W., Cohen, S. 1971. *Gastroenterology* 61:16–24
11. Christensen, J. 1970. *J. Clin. Invest.* 49:681–91
12. Tuch, A., Cohen, S. 1973. *J. Clin. Invest.* 52:14-20
13. Harris, L. D., Winans, C. S., Pope, C. E. 1966. *Gastroenterology* 50:754–60
14. Cohen, S., Harris, L. D. 1971. *N. Engl. J. Med.* 284:1053–56
15. Lipshutz, W., Hughes, W., Cohen, S. 1972. *J. Clin. Invest.* 51:522–29
16. Castell, D. O., Harris, L. D. 1970. *N. Engl. J. Med.* 282:886–89
17. Kline, M. M. et al 1974. *Gastroenterology* 66(4):724 (Abstr.)
18. DiMarino, A. J., Cohen, S. 1973. *J. Clin. Invest.* 52:2264–71
19. Lind, J. F. et al 1969. *Gastroenterology* 56:1078–86
20. Lind, J. F., Crispin, J. S., McIver, D. K. 1968. *Can. J. Physiol. Pharmacol.* 46:233–38
21. Lipshutz, W., Tuch, A., Cohen, S. 1971. *Gastroenterology* 61:454–60
22. Cohen, S., Lipshutz, W. 1971. *J. Clin. Invest.* 50:449–54
23. Lipshutz, W., Cohen, S. 1972. *Am. J. Physiol.* 222:775–81
24. Resin, H. et al 1973. *Gastroenterology* 64:946–49
25. Fisher, R. S., DiMarino, A. J., Cohen, S. 1974. *Clin Res* 22(3):358 (Abstr.)

26. Jennervein, H. M. et al 1973. *Gut* 14:861–64
27. Morris, D. W. et al 1974. *Gastroenterology* 66:(4):750 (Abstr.)
28. Nebel, O. T., Castell, D. O. 1977. *Gut* 14:270–74
29. Nebel, O. T., Castell, D. O. 1973. *J. Appl. Physiol.* 35:6–8
30. Lind, J. F., Warrian, W. G., Wankling, W. J. 1966. *Can. J. Surg.* 9:32–38
31. Burnstock, G. G. et al 1970. *Brit. J. Pharmacol.* 40:668–76
32. Cohen, B. R. 1965. *Gastroenterology* 48:(4)864 (Abstr.)
33. DiMarino, A. J., Cohen, S. 1974. *Gastroenterology* 66:1–6
34. Goyal, R. K., Rattan, S. 1973. *J. Clin. Invest.* 52:337–41
35. Rashid, S. 1971. *J. Pharm. Pharmacol.* 23:456–57
36. Pope, C. E. 1972. *Viewpoints Dig. Dis.* 4:1
37. Battle, W. A., Nyus, L. M., Bombeck, C. T. 1973. *Ann. Surg.* 177:560–65
38. Donner, M. W. et al 1966. *Radiology* 87:220–27
39. Crummy, A. F. 1966. *Radiology* 87:501–4
40. Akdarmar, K. et al 1973. *Gastrointest. Endosc.* 19:172–73
41. Ismail-Beigi, F., Horton, P. F., Pope, C. E. 1970. *Gastroenterology* 58:163–74
42. Schuman, B. M., Rinaldo, J. A. 1966. *Gastrointest. Endosc.* 12:14–16
43. Bernstein, L. M., Baker, L. A. 1958. *Gastroenterology* 34:760–81
44. Benz, L. J. et al 1972. *Gastroenterology* 62:1–5
45. Tuttle, S. G., Grossman, M. I. 1958. *Proc. Soc. Exp. Biol. Med.* 98:225–27
46. Booth, D. J., Kemmerer, W. T., Skinner, D. B. 1968. *Arch. Surg.* 96:731–34
47. Kramer, P., Ingelfinger, F. J. 1951. *Gastroenterology* 19:242–54
48. Hollis, J. B., Castell, D. O. 1972. *Gastroenterology* 63:417–22
49. Hiebert, C. A., Belsey, R. 1961. *J. Thorac. Cardiov. Surg.* 42:352–62
50. Hiebert, C. A. 1970. *Am. J. Surg.* 110:365–71
51. Wolf, B. S. 1961. *J. M. Sinai Hosp. New York* 27:404–16
52. Lipshutz, W. H. et al 1974. *Gastroenterology* 67:423–27
53. Farrell, R. L., Castell, D. O., McGuigan, J. E. 1973. *Clin. Res.* 21(3):512 (Abstr.)
54. Lipshutz, W. H. et al 1973. *N. Engl. J. Med.* 289:182–84
55. Csendes, A., Larrain, A. 1972. *Gastroenterology* 63:19–24
56. Cohen, S., Lipshutz, W. H. 1970. *Ann. Intern. Med.* 72:792–97
57. Castell, D. O., Levine, S. M. 1971. *Ann. Intern. Med.* 74:223–27
58. Higgs, R. H., Smyth, R. D., Castell, D. O. 1974. *Gastroenterology* 66(4):710 (Abstr.)
59. Rogers, A. I., Rothman, S. L., Arostegui, M. 1974. *Clin. Res.* 22(1):23 (Abstr.)
60. Malhotra, S. L. 1970. *Am. J. Dig. Dis.* 15:489–96
61. Stanciu, C., Bennett, J. R. 1974. *Lancet* 1:109–12
62. Skinner, D. B., Belsey, R. H. R. 1967. *J. Thorac. Cardiov. Surg.* 53:33–50
63. Nissen, R. 1961. *Am. J. Dig. Dis.* 6:954–61
64. Hill, L. D. 1967. *Ann. Surg.* 166:681–92
65. Rodnan, G. P., Fennell, R. H. 1962. *J. Am. Med. Assoc.* 180:665–70
66. Stevens, M. D. et al 1964. *N.Engl. J. Med.* 270:1218–22
67. Atkinson, M., Summerling, M. D. 1966. *Gut* 7:402–8
68. Neschis, M. S. et al 1970. *Am. J. Dig. Dis.* 15:443–47
69. Creamer, B., Anderson, H. A., Code, C. F. 1956. *Gastroenterologia* 86: 763–75
70. Cohen, S. et al 1972. *J. Clin. Invest.* 51:2663–68
71. Treacy, W. L. et al 1963. *Ann. Intern. Med.* 59:351–56
72. Willerson, J. T. et al. 1970. *Ann. Intern. Med.* 72:17–27
73. Donald, I. 1964. *Practical Obstetrical Problems.* Chicago: Yearb. Med. 3rd ed.
74. Nagler, R., Spiro, H. M. 1961. *J. Clin. Invest.* 40:954–70
75. Lind, J. F. et al 1968. *Can. Med. Assoc. J.* 98:571–74
76. Laki, K. 1971. *Contractile Proteins and Muscle.* New York: Dekker
77. Mandelstam, P. et al 1969. *Gastroenterology* 56:1–12
78. Winship, D. H. et al 1968. *Gastroenterology* 55:173–78
79. Strawczynski, H., Beck, I. T., McKenna, R. D. 1964. *J. Pediat.* 64:17–23
80. Cohen, S. 1974. *Gastroenterology* 67:252–58
81. Kramer, P., Ingelfinger, F. J. 1949. *Am. J. Med.* 7:168–74
82. Butin, J. W. et al 1953. *Gastroenterology* 23:278–86
83. Creamer, B., Olsen, A. M., Code, C. F. 1957. *Gastroenterology* 33:293–301

84. Cohen, S., Lipshutz, W. 1971. *Gastroenterology* 61:814–20
85. Harris, L. D. 1969. *Advan. Intern. Med.* 15:209–13
86. Code, C. F. et al 1958. *An Atlas of Esophageal Motility in Health and Diseases,* Springfield: Thomas
87. Camara-Lopes, L. H. 1961. *Am. J. Dig. Dis.* 6:742–56
88. Seaman, W. B., Wells, J., Flood, C. A. 1963. *Am. J. Roentgenol.* 90:778–91
89. Herrara, A. F. et al 1970. *Am. J. Dig. Dis.* 15:1073–81
90. Kramer, P., Ingelfinger, F. J. 1951. *Gastroenterology* 19:242–51
91. Hightower, N. C., Olsen, A. M., Moersch, H. J. 1954. *Gastroenterology* 26:592–600
92. Sheperd, J. K., Diamant, N. E. 1972. *Gastroenterology* 63:557–63
93. Kramer, P. et al 1967. *Gut* 8:120–27
94. Castro, L., Grossi, C. A. 1963. *Rev. Goiana Med.* 9:3–19
95. Cohen, S., Lipshutz, W., Hughes, W. 1971. *J. Clin. Invest.* 50:1241–47
96. Cannon, W. B., Rosenblueth, A. 1949. *The Supersensitivity of Denervated Structures. A Law of Denervation.* New York: MacMillan
97. Lendrum, F. C. 1937. *Arch. Intern. Med.* 59:474–80
98. Cassella, R. R. et al 1964. *Ann. Surg.* 160:474–87
99. Hawthorne, H. R., Frobese, A. A., Nemir, P. 1956. *Ann. Surg.* 144:653–63
100. Van Goldsenhoven, G. E. et al 1963. *Gastroenterology* 45:326–34
101. Vantrappen, G. et. al 1963. *Gastroenterology* 45:317–25
102. Nemir, P. et al 1971. *Am. J. Surg.* 121:143–49
103. Campos, J. V., Tafari, W. L. 1973. *Gut* 14:910–19
104. Koberle, F. 1958. *Gastroenterology* 34:460–69
105. Bennett, J. R., Hendrix, T. R. 1970. *Gastroenterology* 59:273–79
106. Fleshler, B. 1967. *Gastroenterology* 52:559–64
107. Roth, H. P., Fleshler, B. 1964. *Ann. Intern. Med.* 61:914–22
108. Stiennon, O. 1968. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* 104:617–24
109. Ismay, G. 1952. *Brit. Med. J.* 2:697–703
110. Belsey, R. 1963. *Postgrad. Med. J.* 39:290–98
111. Cassella, R. R., Ellis, F. H., Brown, A. L. *J. Am. Med. Assoc.* 191:379–82
112. Bettarello, A., Tuttle, S. G., Grossman, M. I. 1960. *Gastroenterology* 39:340–46
113. Farrell, R. L., Roling, G. T., Castell, D. O. 1973. *Am. J. Dig. Dis.* 8:646–50
114. Farrell, R. L., Roling, G. T., Castell, D. O. 1974. *Ann. Intern. Med.* 80:573–76
115. Christensen, J., Daniel, E. E. 1968. *J. Pharmacol. Exp. Ther.* 159:243–253
116. Jacoby, H. I., Broche, D. A. 1967. *Gastroenterology* 52:676–84
117. Stanciu, C., Bennett, J. R. 1973. *Gut* 14:275–79
118. Dilawari, D. B., Misiewicz, J. J. 1973. 14:380–82
119. DiMarino, A. J., Cohen, S. 1974. *Fed. Proc.* 33(3):391 (Abstr.)
120. Cohen, S. 1974. *Proc. Soc. Exp. Biol. Med.* 145:1004–7

Copyright 1975. All rights reserved

EFFECTS OF ANESTHESIA ON INTERMEDIARY METABOLISM

❖7160

Edward A. Brunner, M.D., Ph.D., S. C. Cheng, Ph.D., and M. L. Berman, M.D., Ph.D.
Department of Anesthesia, Northwestern University Medical School, Chicago, Illinois 60611

INTRODUCTION

Some order is beginning to emerge from the confusion concerning metabolic effects of anesthetics which Bunker noted only a dozen years ago (1). Early clinical observations, many lacking in scientific merit, have given way to well designed and executed studies focusing on specific problems. Most reports have been directed to carbohydrate or energy metabolism in the intact organism, liver, or central nervous system. Prior reviews of interest have appeared (1–13).

CARBOHYDRATE METABOLISM

Carbohydrate homeostasis reflects the balance of a complex set of interactions including intestinal absorption, peripheral substrate utilization, the effects of endocrine factors (insulin, glucagon, steroids, growth hormone), gluconeogenesis, glycogenolysis in liver and muscle, lipolysis, and renal clearance of glucose. Anesthetics affect all of these.

Systemic Manifestations

Reynoso first observed glycosuria after ether anesthesia in 1853 (14). Most anesthetics, with the possible exceptions of enflurane and methoxyflurane, are associated with hyperglycemia but not with a rise in plasma insulin level (15). The ratio of plasma insulin/plasma glucose falls (16), indicating that certain anesthetics may partially inhibit insulin release from the pancreas. This is best substantiated for halothane. The role of sympathetic inhibition in this anesthetic depression of insulin secretion has yet to be fully explored, although one observation (17) reports that phentolamine did not prevent the failure of insulin response to infused glucose during halothane anesthesia.

Skeletal Muscle

Theye and his co-workers attempted to define the organ site of the decrease in whole body oxygen consumption which accompanies general anesthesia (18). They showed that oxygen consumption of in situ canine gracilis and gastrocnemius muscles, two markedly dissimilar muscles, decrease equally with increase in halothane concentration (19, 20). This justifies proportional extrapolation to all skeletal muscle, which accounts for 30 to 40% of total body oxygen consumption. Both halothane and methoxyflurane (21) at clinical concentrations inhibit state 3 (ADP present) respiration of rat skeletal muscle mitochondria with glutamate as substrate, but not with succinate. State 4 respiration (ADP absent) with both glutamate and succinate is virtually unaffected. These anesthetics block electron transport in skeletal muscle as they do in brain and liver (see below), and thus reduce oxygen consumption and heat production generated by normal muscle activity. The syndrome of malignant hyperpyrexia has been the subject of a recent symposium (22).

In isolated rat hemidiaphragm (23) ether depresses glucose uptake but does not block insulin stimulation of glucose uptake. It increases lactate formation from glycogen; in the presence of insulin lactate formation is doubled. Normally, hemidiaphragm converts to glycogen 70% of the increased glucose uptake due to insulin. In the presence of ether only 50% is converted to glycogen, the remainder appearing as lactate.

Liver

Jowett & Quastel (24) reported that rat liver respiration was not affected by diethyl ether. More recent studies (25, 26) demonstrate that halothane produces a dose-related reversible depression of oxygen consumption and an increase in glycolysis by rat liver slices. Fink (27), using cultured mouse heteroploid cells, showed that methoxyflurane, halothane, chloroform, and diethyl ether depress oxygen uptake and increase glycolytic rate in order of potency related to the hypnotic activity of the drug. One study (28) of liver slices from anesthetized dogs failed to show changes, probably because of loss of volatile anesthetics from the tissue prior to measurement.

Isolated perfused bovine (29), canine (30), and rat liver (31) have been employed to study anesthetic effects on metabolism. With dog liver, 2% chloroform and 1% methoxyflurane caused glucose release, increased lactate/pyruvate ratio, and loss of potassium and cellular enzymes from the liver. Halothane produced no measurable alterations. In bovine and rat liver chloroform, diethyl ether, halothane, and methoxyflurane (where studied) cause depression of oxygen consumption, lactate utilization, gluconeogenesis, and glycogen synthesis but stimulate release of glucose or lactate into the perfusate. Urea synthesis is depressed and cellular integrity as measured by loss of cellular enzymes and potassium is altered reversibly by chloroform and high concentrations of ether. In fed rats there is a marked decrease in tissue content of both 2-oxoglutarate and citrate and a decrease in calculated NAD/NADH ratio for both cytoplasmic and mitochondrial couples. These findings are consistent with stimulation of glycogenolysis and glycolysis and inhibition of gluconeogenesis and oxidative metabolism. The lack of inhibition of fatty acid oxidation (32) and the decrease in NAD/NADH ratios lead to the speculation of anesthetic inhibition in the region of NADH dehydrogenase as a major site of anesthetic action (33).

In vivo studies of liver utilizing freeze-clamping techniques (34) show that halothane, methoxyflurane, trichloroethylene, and diethyl ether all cause elevated blood and tissue glucose levels and depletion of liver glycogen. In contrast to in vitro studies, halothane caused no lactate accumulation in the liver in vivo. Lactate/pyruvate ratios remain normal even when lactate levels rise, as with trichloroethylene and diethyl ether anesthesia. With halothane, decreases in liver ATP and shift of cytoplasmic and mitochondrial redox couples to a more reduced state do not occur in the intact, as they do in the perfused, liver. Decreases in the liver 2-oxoglutarate levels occur but are less severe (20% of normal in vivo, 50% in vitro) and are accompanied by parallel alterations of glutamate, indicating the presence of intact control mechanisms not functional in the isolated perfused preparation.

Attempts to precisely characterize the basic cellular actions of narcotics have led to demonstrations, first, that urethane and certain alcohols produce a metabolic block of electron transfer (35) and later, that amobarbital and other barbiturates inhibit mitochondrial oxidation of pyruvate and fumarate, but not of succinate (36). This amobarbital inhibition was later extended to all NAD-linked mitochondrial oxidations (37). Cohen and co-workers have shown that halothane in concentrations up to 2% (38–40) reversibly blocks NAD-linked substrate oxidation but does not block succinate oxidation. Only amobarbital-sensitive but not amobarbital-insensitive oxidations are blocked. The site of halothane inhibition occurs in the region of Complex I between NADH dehydrogenase and coenzyme Q (Figure 1), at or near the site(s) of the inhibiting action of rotenone and amobarbital (41–43).

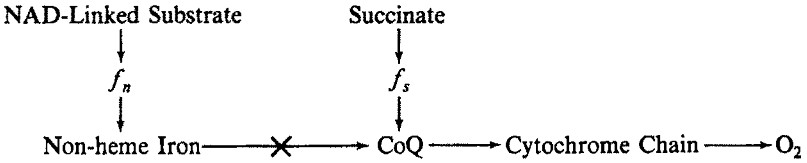


Figure 1. Site of anesthetic inhibition of electron transport.

Methoxyflurane, diethyl ether, enflurane, forane, fluroxene, trichloroethylene, and chloroform in anesthetic concentrations subsequently have been shown to demonstrate a dose-dependent reversible inhibition of electron transport in a way similar to halothane. The concentrations of six anesthetics necessary for 50% inhibition of state 3 respiration bears an inverse log-log relationship to their lipid solubility (44). Potency as electron transport inhibitor in vitro is directly related to anesthetic potency in vivo.

Miller & Hunter (45) suggest that errors in experimental design in some studies have caused inadvertent mitochondrial damage due to exposure to excessive concentrations of halothane and have led to the erroneous conclusion that halothane is a true uncoupling agent. By contrast, studies of mitochondrial calcium transport (45) in which halothane exposure is carefully controlled show that concentrations as high as 4% do not alter the final steady state of calcium accumulation as would a true uncoupler. Moreover, the response of rapid calcium release on addition of a true uncoupler remains intact, indicating that halothane acts differently than a true uncoupler.

There is evidence that methoxyflurane and diethyl ether (41) at anesthetic concentration, and others at higher concentrations (46), have additional actions on mitochondrial function, including inhibition of succinate oxidation, some degree of uncoupling, loss of respiratory control, and alteration of the mitochondrial membrane (46–50). The possibility that anesthetics alter membrane permeability and thereby affect metabolic function by altering the delivery of substrate to the site of utilization has been suggested (41), based on data observed in rat atria and human red cells as well as in mitochondria. Further implications of the above cited findings have yet to be fully evaluated.

Central Nervous System

Metabolism in the nervous system subserves three main functions: 1. to maintain structural integrity, 2. to maintain ionic and electrical gradients across excitable membranes, and 3. to synthesize neurotransmitters. The major metabolic effects of anesthetics relate to the latter two. The following discussion will address effects of volatile anesthetics on energy metabolism, on incorporation and turnover of tracer compounds, and on neurotransmitters.

ENERGY METABOLISM The effect of volatile anesthetics on the respiration of intact brain in vivo is uniformly inhibitory (51–53). The exact extent of this inhibition varies for different agents and species. Several factors (Po_2, Pco_2, blood pressure, vascular resistance, temperature, the anesthetic agent employed, and cerebral blood flow) are important in determining the degree of respiratory inhibition (13, 53). With brain slices, anesthetic agents reduce the additional increment of oxygen utilization due to stimulation by electrical current or high potassium concentration before the resting respiration is altered. The concentration of anesthetic agent required in vitro is usually higher than that in vivo (9). With few exceptions, anesthetic agents depress oxygen consumption by cerebral cortex slices, by chopped brain suspensions, and by brain homogenates (9, 12, 54, 55).

Volatile anesthetic agents cause varied changes in lactate content and glycolysis in the in vivo brain, in slices, and in homogenates (5, 9, 54–59). The recent data demonstrating decreased lactate in brain material recovered with a "brain blower" deserve special mention. Total tissue high energy phosphate content is not altered by volatile anesthetic agents (9, 56, 58–61). Glucose and glycogen can be added to the high energy reserve to represent total energy stores. These increase as metabolic rate is depressed by anesthetics (56–58).

MITOCHONDRIAL AND SYNAPTOSOMAL METABOLISM Crude brain mitochondria can be separated into several subfractions, including synaptosomes and heavy mitochondria. The respiration and oxidative phosphorylation of crude mitochondria supported by pyruvate, glutamate, or 2-oxoglutarate are inhibited by halothane and methoxyflurane but not by nitrous oxide, diethyl ether, chloroform, or cyclopropane (62). Presumably, halothane uncouples oxidative phosphorylation of this crude preparation (63). This idea of "uncoupling," put forth in the early fifties (64), is supported by more recent work using purified mitochondria from rabbit brain and high concentrations of halothane (65), but disputed by work using rat brain mitochondria (66). This latter report demonstrates that halothane inhibits ADP-stimulated respiration of mitochondria metabolizing glutamate. This inhibiting effect is virtually absent with succinate. Calcium uptake is also found to be inhibited by halothane. These results have been confirmed in a concurrent study using both mitochondrial and synaptosomal preparations (67). The succinate reversal of anesthetic-inhibited respiration was first observed many years ago (55, 68) without adequate explanation. It appears likely that further studies will confirm an anesthetic sensitive site in electron transport similar to that already identified in liver mitochondria.

TRACER EXPERIMENTS The validity of results based on the cerebral metabolism of labeled substrates is often in question due to the large cerebral blood flow which introduces a "contaminating" effect from metabolism of other body tissues. This compounds difficulties caused by the postmortem changes which are well known in neurochemistry. Anesthetic agents alter amino acid labeling in the brain and spinal cord. In general (56–58, 60, 69–74), the incorporation of glucose carbon into aspartate, glutamate, glutamine, and γ-aminobutyrate is reduced but similar incorporation into alanine is variable. Glutamate content is unchanged or decreased. Incorporation of carbons from acetate, butyrate, and citrate is not affected. These differences observed in the incorporation of labeled carbon from different substrates can be interpreted by the concept of metabolic compartmentation (75). Two groups of substrates are known to be metabolized by different compartments and their correlated anatomic entities have been respectively assigned to nerve endings and glia. The anesthetic agents appear to inhibit primarily the glucose utilizing compartment, i.e. the nerve endings (69).

NEUROTRANSMITTERS This section includes data on the effect of amobarbital, chloralose, allobarbital, hydroxydione, methylpentyol, pentobarbital, thiopental, phenobarbital, and the volatile anesthetics on synthesis, degradation, release, uptake, and other mechanisms of removal from site(s) of action of neurotransmitters. Most inhibit the synthesis of acetylcholine (ACh) (12, 68, 69, 76–78) and cause increased ACh content in vivo (8, 79, 80), which is attributed to nonutilization. In contrast, the ACh content of slices exposed to halothane is lower than that of controls and its synthesis from acetate is not altered (69). The release of ACh is reduced by anesthetic agents (11, 76, 81) and its turnover rate is decreased (69, 78).

Anesthetic agents cause reduced synthesis of γ-aminobutyrate (GABA) in rat cortex slices (69); release is not affected (81). In stimulated slices barbital causes increased GABA release (82). GABA content in rat brain varied (83, 84) and GABA uptake by rat cortex slices also varied (69, 82, 85).

The primary manner of the removal of GABA from synaptic areas is by re-uptake by cells near the synapses followed by enzymatic degradation via the GABA shunt, a unique metabolic pathway in the brain. A glial function for such removal has been advanced. This whole matter concerning GABA requires further clarification.

HYPOTHESIS OF ANESTHETIC ACTION The increase in energy stores, the reduced metabolic rate (oxygen consumption), and the reduced incorporation of glucose carbon into intermediate metabolites suggest a decreased utilization of energy in brain during anesthesia. Earlier theories based on the inhibition of energy production as the primary cause of anesthesia failed when the energy stores were found not to decrease and the reduction of cerebral NAD could not be demonstrated (59, 86). In vivo inhibition of synaptic transmission was found to precede axonal block by anesthetic agents, suggesting that neurotransmitters may play a crucial role in anesthesia mechanisms. Two lines of thought have been proposed. One suggests that the increased cytoplasmic calcium concentration, induced by impaired mitochondrial uptake of calcium under influence of anesthetic agents, may cause reduced sensitivity to excitation by ACh (66, 67). This idea requires no participation of an inhibitory synapse which some physiological experiments have suggested. The other theory postulates that degradation of GABA is retarded by anesthetic block of Complex I. This causes a rise in synaptic GABA content followed by synaptic inhibition. Regardless of the cause of the decreased synaptic activity, the outward manifestation would be anesthesia. A decrease in synaptic work requires less utilization of energy for ionic transport. Extra energy is shifted to the synthesis of energy-rich compounds such as glycogen, lipids, and glutamine. The energy-producing reactions (respiration and glycolysis) are somewhat reduced. This oversimplified hypothesis is attractive. It agrees with most neurochemical findings. The details are more difficult to formulate. For example, glial functions, which are just being investigated, may contribute heavily in the overall mechanism of anesthesia.

Miscellaneous

The effects of inhalation anesthetics on myocardial metabolism have been included in a recent review of Merin (87) and interested readers are referred thereto. Isolated studies of anesthetic effects on lung and kidney metabolism and of adrenal enzyme activity have been reported. A few isolated studies of anesthetic effect on specific enzymes have also been reported, but systematic analysis of results is not yet possible.

FAT METABOLISM

Studies of anesthetic effects on lipid metabolism in man have focused on the changes in the plasma levels of nonesterified fatty acids (NEFA). In most studies adequate control of pertinent parameters is lacking or not recorded. Pre-anesthetic emotional stress and surgical stress are more significant than nitrous oxide-halothane anesthesia in causing a marked rise in plasma NEFA (17). Cyclopropane causes marked elevation of plasma NEFA which can be inhibited by prior β-adrenergic blockade or reversed by glucose loading. Halothane-nitrous oxide anesthesia causes a rise in plasma NEFA but not as great as with cyclopropane; diethyl ether causes no elevation in plasma NEFA (88). Halothane and methoxyflurane cause a significant rise in plasma NEFA (10). Patients anesthetized by four combinations of anesthetics, halothane, nitrous oxide, thiopental, and droperidol with fentanyl, for renal arteriography showed no significant changes in plasma NEFA over pre-anesthetic levels, suggesting that in absence of a surgical stress these anesthetics fail to affect lipid metabolism (89). General anesthesia with ether and neuroleptic agents and spinal anesthesia cause no significant alterations in NEFA (15). The lack of change of plasma NEFA following spinal anesthesia with subsequent sympathetic blockade suggests that elevated plasma NEFA with general anesthetics may involve sympathetic nervous system stimulation.

Nonclinical studies on the effects of inhalational anesthetics on lipid metabolism are sparse. The uptake of free fatty acids by the liver of ether-anesthetized rats is markedly increased compared to unanesthetized rats (90), but this may be related to circulatory changes. Exposure of rat hepatoma cells to halothane results in a marked increase in lipid synthesis (neutral and phospholipid fractions), a slight reduction in DNA and protein synthesis, and no effect on RNA synthesis. Acetate uptake is increased in the presence of halothane (91).

PROTEIN AND NUCLEIC ACIDS METABOLISM

Studies of anesthesia and protein metabolism are scarce. Negative nitrogen balance persists longer following ether than after spinal anesthesia (92). Ether and trichloroethylene, but not nitrous oxide, significantly decrease the rate of incorporation of L-leucine into the exocrine cells of the rat pancreas during 15 min of anesthesia. Inhibition also occurs in vitro when slices of pancreas are exposed to anesthetic gases (93). Halothane depresses uptake of leucine and thymidine by rat hepatoma cells with subsequent reduction of DNA and protein synthesis and no effect on RNA synthesis (91). Sulfhydryl groups in brain homogenates suspended in solutions of diethyl ether, chloroform, halothane, methoxyflurane, and trichloroëthylene are decreased (94). A correlation exists between anesthetic potency and sulfhydryl reducing effect, suggesting the possibility of conformational changes in brain protein.

Hinkley & Telser (95) have recently summarized studies of effects of general anesthetics on the microtubule-microfilament systems which have been implicated in cytoplasmic movement. Allison & Nunn (96) suggest that reversible depolymerization of neuronal microtubules by anesthetics might be the basis of general anesthesia, but Sauberman & Gallagher (97) have shown narcosis can occur without microtubular disruption. Actin-like microfilaments are reversibly depolymerized with clinical concentrations of inhalation anesthetic agents (98). These filaments in the CNS are thought to involve neurotransmitter release and interaction of anesthetic and microfilament protein has been postulated as a mechanism by which volatile anesthetics affect biomotion phenomena (95).

SUMMARY

Major inhalational anesthetics cause inhibition in the electron transport chain in the region of Complex I resulting in decreased oxygen utilization, inhibition of metabolism of NAD-linked substrates, but not of succinate, inhibition of mitochondrial calcium uptake, and depression of synaptic transmission because of postulated changes in ACh sensitivity or GABA inhibition. Many cellular metabolic effects in CNS and other tissues are secondary to the above. Many metabolic changes noted with anesthetics occur subsequent to activation of the sympathetic nervous system either directly by the anesthetic or by surgical stimulation in the presence of light anesthesia. Many important studies remain to be done.

Literature Cited

1. Bunker, J. P. 1963. Neuroendocrine and other effects on carbohydrate metabolism during anesthesia. *Anesthesiology* 24:515–23
2. Bunker, J. P. 1962. Metabolic acidosis during anesthesia and surgery. *Anesthesiology* 23:107–22
3. Bunker, J. P., Vandam, L. D. 1965. Effects of anesthesia on metabolism and cellular functions. *Pharmacol. Rev.* 17:183–263
4. Fink, B. R. 1969. Cellular biology of anesthesia. *Clin. Anesth.* 3:109–30
5. Fink, B. R., Haschke, R. H. 1973. Anesthetic effects on cerebral metabolism. *Anesthesiology* 39:199–215
6. Greene, N. M. 1963. *Inhalation Anesthetics and Carbohydrate Metabolism.* Baltimore: Williams & Wilkins. 143 pp.
7. Greene, N. M. 1968. Halothane and metabolism. *Clin. Anesth.* 1:181–95
8. Kitahata, L. M., Schmidt, K. F. 1968. Effects of halothane on neuronal tissue. *Clin. Anesth.* 1:61–84
9. McIlwain, H. 1966. *Biochemistry and the Central Nervous System.* Chap 13, 338–58. Boston: Little, Brown
10. Merin, R. G., Samuelson, P. N., Schalch, D. S. 1971. Major inhalation anesthetics and carbohydrate metabolism. *Anesth. Analg.* 50:625–32
11. Pepeu, G. 1974. *Progr. Neurobiol.* 2:257–88
12. Quastel, J. H. 1965. Effects of drugs on metabolism of the brain in vitro. *Brit. Med. Bull.* 21:49–56
13. Smith, A. L., Wollman, H. 1972. Cerebral blood flow and metabolism. *Anesthesiology* 36:378–400
14. Reynoso, A. 1853. Note sur le passage du sucre dans les urines a propas d'une note due Docteur Harley sur le meme suject. *C. R. Soc. Biol.* 5:116–19
15. Oyama, T. 1973. *Anesthetic Management of Endocrine Disease.* Chap 9, 118–39, 199–207. New York: Springer-Verlag
16. Greene, N. M. 1974. Insulin and anesthesia. *Anesthesiology* 41:75–79
17. Allison, S. P., Tomlin, P. J., Chamberlain, M. J. 1969. Some effects of anaesthesia and surgery on carbohydrate and fat metabolism. *Brit. J. Anaesth.* 41:588–93
18. Braun, U., Hensel, I., Kettler, D., Lohr, B. 1971. Influence of methoxyflurane, halothane, dipiritramide, barbiturate and ketamine on the total oxygen consumpation of the dog. *Anaesthesist* 20:369–75
19. Theye, R. A. 1970. Effect of halothane on canine gastrocnemius-muscle oxygen consumption. *Anesth. Analg.* 49:680–86
20. Theye, R. A., Brown, A. L. Jr., Van Dyke, R. A. 1971. The effects of halothane and succinylcholine on oxygen uptake of the canine gracilis muscle. *Anesthesiology* 35:304–8
21. Britt, B. A., Kalow, W., Endrenyi, L. 1972. Effects of halothane and methoxyflurane on rat skeletal muscle mitochondria. *Biochem. Pharmacol.* 21:1159–69
22. Gorden, R. A., Britt, B. A., Kalow, W., Eds. 1973. *International Symposium on Malignant Hyperthermia.* Springfield: Thomas
23. Brunner, E. A. 1969. The effects of diethyl ether on carbohydrate metabolism in skeletal muscle. *Anesthesiology* 30:24–28
24. Jowett, M., Quastel, J. H. 1937. The effect of ether on brain oxidations. *Biochem. J.* 31:1101–16

25. Hoech, G. P., Matteo, R. S., Fink, B. R. 1966. Effect of halothane on oxygen consumption of rat brain, liver and heart and anaerobic glycolysis of rat brain. *Anesthesiology* 27:770–77
26. Besombe, J., Longmuir, I. S. 1971. Effect of halothane on oxygen consumption of rat liver slices. *Brit. J. Anaesth.* 43:1036–42
27. Fink, B. R., Kenny, G. E., Simpson, W. E. 1969. Depression of oxygen uptake in cell culture by volatile, barbiturate and local anesthetics. *Anesthesiology* 30:150–55
28. MacBeth, R. A., Haley, F. C. H., Bekesi, J. G. 1961. Cellular metabolism of surviving tissue slices following various types of anesthesia. *Surg. Forum* 12:11–12
29. Bombeck, C. T., Aoki, T., Smuckler, E. A., Nyhus, L. M. 1969. Effects of halothane, ether and chloroform on the isolated, perfused, bovine liver. *Am. J. Surg.* 117:91–107
30. Strunin, L. et al 1970. A comparison of the effects of halothane and methoxyflurane on the isolated perfused canine liver. *Acta Anaesthesiol. Scand. Suppl.* 37:203–4
31. Biebuyck, J. F., Lund, P., Krebs, H. A. 1972. The effects of halothane (2-bromo-2-chloro-1,1,1-trifluorethane) on glycolysis and biosynthetic processes of the isolated perfused rat liver. *Biochem. J.* 128:711–20
32. Biebuyck, J. F., Lund, P., Krebs, H. A. 1972. The protective effect of oleate on metabolic changes. *Biochem. J.* 128:721–23
33. Biebuyck, J. F. 1973. Anesthesia and hepatic metabolism. *Anesthesiology* 39:188–98
34. Biebuyck, J. F., Lund, P. 1974. Effects of halothane and other anesthetic agents on the concentrations of rat liver metabolites in vivo. *Mol. Pharmacol.* 10:474–83
35. Keilin, D., Hartree, E. F. 1940. Succinic dehydrogenase-cytochrome system of cells: intracellular respiratory system catalysing aerobic oxidation of succinic acid. *Proc. Roy. Soc. London B* 129:277–306
36. Eiler, J. J., McEwan, W. K. 1949. The effect of pentobarbital on aerobic phosphorylation in brain hemogenates. *Arch. Biochem. Biophys.* 20:163–65
37. Ernster, L., Jalling, O., Low, H., Lindberg, O. 1955. Alternative pathways of mitochondrial DPNH oxidation, studied with amytal. *Exp. Cell. Res. Suppl.* 3:124–32
38. Cohen, P. J., Marshall, B. E., Harris, J. E., Lecky, J. H., Rosner, B. S. 1968. *Fed. Proc.* 27:2744 (Abstr.)
39. Cohen, P. J., Marshall, B. E. 1968. *Toxicity of Anesthetics,* ed. B. R. Fink, Chap. 3, 24–36. Baltimore: Williams & Wilkins
40. Cohen, P. J., Marshall, B. E., Lecky, B. A. 1969. Effects of halothane on mitochondrial oxygen uptake: site of action. *Anesthesiology* 30:337
41. Cohen, P. J., McIntyre, R. A. 1972. *Cellular Biology and Toxicity of Anesthetics,* ed. B. R. Fink, Chap 10, 109–16. Baltimore: Williams & Wilkins. 328 pp.
42. Hall, G. M., Kirtland, S. J., Baum, H. 1973. The inhibition of mitochondrial respiration by inhalational anaesthetic agents. *Brit. J. Anaesth.* 45:1005–9
43. Harris, R. A., Munroe, J., Farmer, B., Kim, K. C., Jenkins, P. 1971. Action of halothane upon mitochondrial respiration. *Arch. Biochem. Biophys.* 142: 435–44
44. Nahrwold, M. L., Cohen, P. J. 1973. The effects of forane and fluroxene on mitochondrial respiration. *Anesthesiology* 39:437–44
45. Miller, R. N., Hunter, F. E. 1971. Is halothane a true uncoupler of oxidative phosphorylation? *Anesthesiology* 35: 256–61
46. Miller, R. N., Hunter, F. E. 1970. The effect of halothane on electron transport, oxidative phosphorylation and swelling in rat liver mitochondria. *Mol. Pharmacol.* 6:67–77
47. See Ref. 41, 93–108
48. Miller, R. N., Schumer, W., Erve, P. R. 1972. The effects of halogenated anesthetics on mitochondrial function. *Anesthesiology* 36:625–27
49. Schumer, W., Erve, P. R., Obernolte, R. P., Bombeck, C. T., Sadove, M. S. 1971. The effect of inhalation of halogenated anesthetics on rat liver mitochondrial function. *Anesthesiology* 35: 253–55
50. Schumer, W., Erve, P. R., Obernolte, R. P., Bombeck, C. T., Sadove, M. S. 1972. Halogenated anesthetics' effect on rat liver mitochondria. *Z. Exp. Chir.* 5:13–16
51. Michenfelder, J. D., Theye, R. A. 1972. Effects of cyclopropane on canine cerebral blood flow and metabolism. *Anesthesiology* 37:32–39
52. Michenfelder, J. D., Theye, R. A. 1973. Effects of methoxyflurane on canine

cerebral metabolism and blood flow. *Anesthesiology* 38:123–27

53. Smith, A. L. 1973. The mechanism of cerebral vasodilation by halothane. *Anesthesiology* 39:581–87
54. Matteo, R. S., Hoech, G. P., Hoskin, F. C. G. 1969. Effects of cyclopropane and diethyl ether on tissue oxygen consumption and anaerobic glycolysis of brain in vitro. *Anesthesiology* 30:156–63
55. Hoesin, E. A., Stachiewicz, E., Bourne, W., Demstedt, O. F. 1955. The influence of nitrous oxide on the metabolic activity of brain tissue. *Anesthesiology* 16:708–15
56. Biebuyck, J. F., Hawkins, R. A. 1972. The effect of anaesthetic agents on brain tissue metabolite patterns. *Brit. J. Anaesth.* 44:226–27
57. Brunner, E. A., Passonneau, J. V., Molstad, C. 1971. The effect of volatile anaesthetics on levels of metabolites and on metabolic rate in brain. *J. Neurochem.* 18:2301–16
58. Heim, F., Esteler, C.-J., Ammon, H. P. T., Picker, D. 1966. Effect of trichloroethylene on brain levels of high energy phosphate, coenzyme A and glycolytic metabolites in white mice. *Arch. Int. Pharmacodyn. Ther.* 162:311–18
59. Nilsson, L., Siesjo, B. K. 1970. The effect of anesthetics upon labile phosphates and upon extra- and intra-cellular lactate, pyruvate and bicarbonate concentrations in the rat brain. *Acta Physiol. Scand.* 80:235–48
60. Goldberg, N. D., Passonneau, J. V., Lowry, O. H. 1966. The effects of changes in brain metabolism on the levels of citric acid cycle intermediates. *J. Biol. Chem.* 17:3997–4003
61. Doring, H. J., Olbrisch, R. R., Schrader, J., Lang, H. 1970. Electrocorticograms, steady potential of the brain and energy-rich phosphate fractions of the cerebral cortex during anesthetic overdosage, ischemia and cyanide poisoning. *Pfleuger S Arch.* 319:12–35
62. Gatz, E. E., Schoettger, J. D., Morgan, J. R., Bird, J. S., Jones, J. R. 1971. *Fed. Proc.* 30:442 (Abstr. 1378)
63. Gatz, E. E., Jones, J. R. 1969. *Fed. Proc.* 28:356 (Abstr. 353)
64. Brody, T. M., Bain, J. A. 1954. Barbiturates and oxidative phosphorylation. *J. Pharmacol. Exp. Ther.* 110:148–56
65. Smith, H. 1973. Effect of halothane upon rabbit brain mitochondria. *Biochem. Pharmacol.* 22:773–81
66. Rosenberg, H., Haugaard, N. 1973. The effects of halothane on metabolism and calcium uptake in mitochondria of the rat liver and brain. *Anesthesiology* 39:44–53
67. Miller, R. N., Miller, E., Caine, J., Stabler, M. 1974. Synaptosomes, a possible system for the study of synaptic transmission. *Anesth. Analg.* 53:132–41
68. Johnson, W. J., Quastel, J. H. 1953. A comparison of the effects of narcotics and 2,4-dinitrophenol on sulfanilamide acetylation. *J. Biol. Chem.* 205:163–71
69. Cheng, S.-C., Brunner, E. A. 1975. *Molecular Mechanisms of Anesthesia*, ed. B. R. Fink. New York: Raven. In press
70. Cremer, J. E., Lucas, H. M. 1971. Sodium pentobarbitone and metabolic compartments in rat brain. *Brain Res.* 35:619–21
71. Hakim, A. M., Moss, G. 1974. The effect of ether anesthesia on cerebral glucose metabolism—the pentose phosphate pathway. *Anesthesiology* 40: 261–67
72. Hassler, C., Hassler, R., Okada, Y., Bak, I. J. 1971. Pre-ictal and ictal changes of serotonin, GABA and glutamate contents in different regions of rabbit brain during methoxypyridoxine-induced seizures. *Acta Neurol. Latinoamer.* 17:595–611
73. Reynolds, A. P., Gallagher, B. B. 1973. The effect of hexafluorodiethyl ether (flurothyl) on the metabolism of rat brain amino acids labelled by (U-^{14}C) glucose. *Life Sci.* 13:87–95
74. Reynolds, A. P., Watkins, J. C. 1972. The effect of strychnine and of electrical stimulation on the labelling of γ-aminobutyric acid and other free amino acids from (U-^{14}C) glucose in the spinal cord of the nembutalized rat. *Brain Res.* 36:343–51
75. Balazs, R., Cremer, J. E., 1973. *Metabolic Compartmentation in the Brain.* London: Macmillan. 383 pp.
76. Grewaal, D. S., Quastel, J. H. 1973. Control of synthesis and release of radioactive acetylcholine in brain slices from the rat. *Biochem. J.* 132:1–14
77. Sharkawi, M. 1970. Effects of morphine and pentobarbitone on acetylcholine synthesis by rat cerebral cortex. *Brit. J. Pharmacol.* 40:86–91
78. Sundwall, A. 1973. Effect of pentobarbital on the turnover of acetylcholine in brain nerve terminals in vivo. *Brain Res.* 62:531–36
79. Domino, E. F., Wilson, A. E. 1972. Psychotropic drug influences on brain ace-

tylcholine utilization. *Psychopharmacologia* 25:291–98
80. Jones, B. E., Guyenet, P., Cheramy, A., Gauchy, C., Glowinski, J. 1973 The in vivo release of acetylcholine from cat caudate nucleus after pharmacological and surgical manipulations of dopaminergic nigrostriatal neurons. *Brain Res.* 64:355–69
81. Tappaz, P. M., Tacheco, H. 1973. Effects of convulsant and anticonvulsant drugs on the spontaneous and induced release of GABA(^{14}C) from cortex slices of rat brain. *J. Pharmacol. (Paris)* 4:433–52
82. Cutler, R. W. P., Dudzinski, D. S. 1974. Effect of pentobarbital on uptake and release of (^{3}H)GABA and (^{14}C) glutamate by brain slices. *Brain Res.* 67:546–48
83. Blagoeva, P., Longo, V. G., Masi, I., Scotti de Carolis, A. 1972. Alterations in behavior, EEG, and brain cortical GABA content following prolonged administration and subsequent withdrawal of barbital in the rat. *Behav. Biol.* 7:755–60
84. Tsuji, H., Balagot, R. C., Sadove, M. S. 1963. Effect of anesthetics on brain gamma-aminobutyric and glutamic acid levels. *J. Am. Med. Assoc.* 183:659–61
85. Harris, M., Hopkin, J. M., Neal, M. J. 1973. Effect of centrally acting drugs on the uptake of γ-aminobutyric acid (GABA) by slices of rat cerebral cortex. *Brit. J. Pharmacol.* 47:229–39
86. Chance, B., Cohen, P., Jobsis, F., Schoener, B. 1962. Intracellular oxidation-reduction states in vivo. *Science* 137:499–508
87. Merin, R. G. 1975. See Ref. 69
88. Cooperman, L. H. 1970. Plasma free fatty acid levels during general anesthesia and operation in man. *Brit. J. Anaesth.* 42:131–35
89. Tarhan, S., Fulton, R. E., Moffitt, E. A. 1971. Body metabolism during general anesthesia without superimposed surgical stress. *Anesth. Analg.* 50:915–23
90. Schotz, M. C., Olivecrona, T. 1966. The effect of anesthesia on the fate of injected free fatty acid. *Biochim. Biophys. Acta* 125:174–75
91. Ishii, D. N., Corbascio, A. N. 1971. Some metabolic effects of halothane on mammalian tissue culture cells in vitro. *Anesthesiology* 35:427–38
92. Manku, R. S. 1970. Effect of anesthesia on protein metabolism in patients undergoing hernioplasty. *Anesth. Analg.* 49:446–49
93. Kramer, M. F., Poort, C. 1972. Influence of anaesthetic drugs on amino acid incorporation in the rat pancreas. *Biochem. Pharmacol.* 21:441–47
94. Ho, I. K., Keats, A. S. 1969. Effect of anesthetics on sulfhydryl content of rat brain. *Pharmacology* 2:257–64
95. Hinkley, R. E., Telser, A. G. 1974. See Ref. 69
96. Allison, A. C., Nunn, J. F. 1968. Effects of general anesthetics on microtubules: a possible mechanism of anesthesia. *Lancet* 3:1326–29
97. Sauberman, A. J., Gallagher, M. L. 1973. Mechanisms of general anesthesia: failure of pentobarbital and halothane to depolymerize microtubules in mouse optic nerve. *Anesthesiology* 38:25–31
98. Wiklund, R. A., Allison, A. C. 1972. Effects of anaesthetics on the motility of Dictyostelium discoideum. *Nature (New Biol.)* 239:221–22

Copyright 1975. All rights reserved

GONADOTROPIN-RELEASING HORMONE

◆7161

Samuel S. C. Yen, M.D. [1]

Department of Reproductive Medicine, University of California, San Diego, La Jolla, California 92037

INTRODUCTION

The existence of a hypothalamic gonadotropin-releasing hormone long suspected from indirect evidence (1) has now been conclusively proven. In 1971, the primary structure of porcine (2) and subsequently ovine (3) luteinizing hormone-releasing hormone or factor (LRH or LRF) was determined to be the decapeptide *p*Glu-His-Trp-Ser-Tyr-Gly-Leu-Arg-pro-Gly-NH_2. Like that of native LRF, its synthetic replicate has been shown to modify not only the secretion of luteinizing hormone (LH) but that of follicle stimulating hormone (FSH) as well. With the availability of this synthetic peptide, enormous amounts of new information have been generated in both experimental and clinical situations. While the data have provided new insights to the physiology and pathophysiology of the hypothalamic-pituitary-gonadal axis, several findings have created unexpected, challenging, and renewed concepts. This article will briefly review the current status of LRF physiology and clinical effects. Other recent reviews are available (4–6). The term LRF, which is synonymous to gonadotropin-releasing hormone (Gn-RH) and luteinizing hormone-releasing hormone (LH-RH), will be used in this review.

BIOSYNTHESIS AND THE SITE OF PRODUCTION OF LRF

There has been no definitive study reported concerning the biosynthesis of LRF at the time of this review. Preliminary experiments (7) indicate that ^{3}H-Arg is incorporated during in vitro incubation of fragments of the rat hypothalamus in the peptide that has the same mobility as a (^{3}H)-LRF decapeptide. Elucidation of the biosynthesis and control of LRF is undoubtedly forthcoming and will add, as did our knowledge of thyrotropin-releasing factor, a new dimension in the understanding of the feedback regulation of the hypothalamic-pituitary unit.

---

[1] Supported in part by Rockefeller Foundation (RF-70029) and by Agency for International Development/CSD 2785.

Present knowledge concerning the site or sites of origin of LRF in the hypothalamic hypophysiotropic area is limited. Although some evidence may indirectly be derived from results of experiments using hypothalamic lesions, such studies have the common drawback of not allowing a clear distinction between the destruction of an LRF-producing area and the interruption of a neural pathway which either leads to the site of release or conveys LRF from the site of production. Similar difficulties in interpretation also apply to data obtained with electrical stimulation of hypothalamic structures.

The recent availability of a specific antiserum to LRF (8, 9) has afforded a sensitive tool to study the site(s) of synthesis of this peptide and the pathway for its transport and secretion into the hypophyseal portal system. Initially, by using fluorescent antibody studies, the immunoreactive LRF was found in the median eminence in processes near portal capillaries and in perikarya of neurons scattered throughout the medial basal hypothalamus of ovariectomized animals (10). Most recently however, using the immunoperoxidase method by which autofluorescense can be avoided, a clearer definition has been revealed (11); in the hypothalamus of the mouse, LRF was localized in the cytoplasm of tanycytes in the median eminence. Large amounts of granular reaction products were seen in many of these specialized ependymal cells: in their cell body lining of the third ventricle, throughout their multiple branching processes, and in their terminals on portal capillaries of the zona externa. The arcuate nucleus of both sexes was the only other hypothalamic region in which immunoreactive LRF was found. Reaction products were limited to the perikarya of the neuron. Most important is the failure to detect LRF along the axons known to run from the cell bodies of the arcuate neurons to the portal bed. These preliminary studies do not provide any direct clue to the functional significance of this distribution, although they do suggest that LRF may be synthesized in arcuate neurons and that the larger amount of LRF in tanycytes may represent storage and the site of concentration, in addition to transport, for later secretion into the portal system on demand. The release of tanycyte LRF might be triggered by the direct effect of gonadal steroids or catecholamines, which are known to have profound effects on the subcellular morphology of tanycytes, or by dopaminergic neurons, which terminate on them (12).

REGULATION OF LRF SECRETION

Until recently the control and modulation of LRF secretion has been inferred primarily from indirect measurements of circulating gonadotropin levels in response to various physiological manipulations. Preliminary studies have been carried out, however, in which biologically active LRF has been detected directly in the rat hypophyseal portal system (13). In addition, superfusion of the hemipituitary with portal blood in rats has indicated that castration causes an increase in hypophyseal portal LRF concentration and very low LRF activity in portal blood during the diestrus stage of the ovulatory cycle (14). Measurements of LRF activity by bioassay and by recently developed radioimmunoassay in peripheral plasma in humans has provided the initial insight that gonadal steroids exert both negative and positive feedback on hypothalamic LRF secretion; bioassayable LRF concentrations were significantly elevated in plasma of postmenopausal women as compared with the plasma levels obtained in premenopausal women (15), indicating an inhibitory action of LRF secretion by gonadal steroids. Preliminary data suggest that estrogen may promote LRF secretion; increased circulating bioassayable and radioimmunoassayable LRF activities at midcycle in humans have been reported (16, 17); LRF activity in plasma has been found to be increased within 24 hr of estrogen administration (18); the elevated plasma LRF activity found in long term castrated males showed a further increase after long term estrogen treatment (19). Although these findings support the long postulated feedback regulation between estrogen and hypothalamic LRF release, implication of this relationship in a physiological content remains to be established. Moreover, questions such as the effect of estrogen on the metabolism and degradation of LRF within the circulation and pituitary must be answered before a meaningful correlation can be made.

In spite of many valuable accomplishments during the past few years, the neurons responsible for the synthesis and the actual modus operandi in the secretion of LRF remain speculative at the present time. The situation is further complicated by the fact that catecholamines are abundant in the hypothalamus and seemingly serve as neurotransmitters, representing a critical step in the regulation of hypothalamic hormones (20). The catecholamines, dopamine and noradrenalin, are present in distinct neuronal pathways in the brain; there is a prominent dopaminergic system consisting of small neurons located in the arcuate necleus and axons projecting to the external layer of the median enimence (20, 21). The nondopaminergic neurons in the same region are currently thought to represent the releasing factor secreting neurons (see above). Preliminary studies have also revealed the presence of noradrenalin together with dopamine in the basal hypothalamus. Semiquantitative studies have shown fluctuation of catecholamine content related to sex steroid environment and castration (20). The mechanism of interaction between catecholamine (adrenergic axons) and hypophysiotropic hormones (peptidergic axons) has been proposed to be the axo-axonic synapsis between these two separate populations of neuronal elements, but the existence of such an anatomical connection has never been established. Alternatively, a simple mechanism for adrenergic modulation of LRF neurons is the control of local blood flow, thereby increasing permeability of the portal vascular system to permit the entry of increased amounts of LRF. A fascinating possibility for the role of catecholamines in the control of ependymal cells of the median eminence has recently been speculated (12, 22, 23); they may modulate the function of these ependymal cells, the tanycytes, which comprise the median eminence connecting the floor of the third ventricle with the portal vessels. These cells possess microvilli on the surface facing the ventricle and have abundant microtubules and microfilaments on the surface facing the wall of the fenestrated portal vessels. A transport mechanism appears to exist, since protein-size molecules can be recovered in the portal vessels after they are injected into the third ventricle (22). Information in this direction is forthcoming and may prove to be a landmark in the understanding of the regulation of neurosecretory mechanics, whereby neurons are able to talk to one another by a local messenger (neurotransmitters). There is good evidence to support this speculation; injection of dopamine into the third ventricle but not into the general circulation leads to a rapid increase in LRF and prolactin inhibitory factor (PIF) activities in the portal blood (13), suggesting dopamine-mediated LRF and PIF regulation.

Since it is now established that pituitary gonadotropin release is pulsatile in character (24), a similar time-related pulsatile discharge of LRF has been postulated (24). Although preliminary, Seyler & Reichlin have presented data which suggest episodic fluctuations of LRF in the plasma of hypogonadal women but which did not correlate with LH spikes (25).

METABOLISM OF LRF

The availability of isotopically labeled forms of LRF and the development of a specific radioimmunoassay for LRF have afforded tools for examining the metabolic fate of LRF. The clearance of exogenous ^{3}H-LRF (26) and ^{125}I-LRF (27) from plasma in rats as well as in humans is exceedingly rapid with an initial half-life (T 1/2) of 4 min. The initial distribution space approximated plasma volume (26, 27). A similar initial disappearance rate of exogenous LRF was found by radioimmunoassay (9, 28). The disappearance curve in all of these studies appears to follow at least two linear components: an initial fast component (T 1/2 = 4 min) and a second slower component with a T 1/2 of approximately 50 min. One hour after the injection of the labeled LRF, 48% of the dose of radioactivity and 73% after 24 hr was found in the urine. No biologically active LRF was found in the urine. Since 80–90% of the radioactivity in the urine was found to be associated with a dipeptide pyroglutamyl-histidine, it was suggested that one of the mechanisms for inactivation of LRF is the cleavage of this dipeptide moiety from the N-terminus of LRF (26). Tissue accumulation of LRF after the injection of labeled hormone was found in the anterior pituitary gland and the pineal gland as well as in the kidney and liver (27). One hour after the injection of ^{3}H-LRF, there was a marked accumulation of radioactivity, as indicated by a tissue to plasma (*T/P*) ratio of 21.9 for kidney, 15.1 for posterior pituitary, 11.6 for pineal gland, 2.5 for liver, and 2.1 for anterior pituitary. The high *T/P* ratio for the liver and kidney indicates that these organs may be the primary sites for the inactivation of LRF. While the increased *T/P* ratio is expected to occur in the anterior pituitary, the high *T/P* ratio observed in the posterior pituitary and more particularly in the pineal is surprising. The latter has recently been shown to contain LRF in greater amounts than the hypothalamus with a ratio of 20:1 (29). Whether this high concentration of LRF in the pineal represents accumulated and/or synthesized LRF de novo is of primary importance in the implication of the pineal gland in the regulatory role of hypothalamic LRF and the reproductive function. It is evident that much excitement is generated by the possible role of cerebrospinal fluid (CSF), pineal gland, and tanycytes as new pathways in the regulation of the adenohypophysis.

MECHANISM OF ACTION (STRUCTURE/FUNCTION RELATIONSHIP)

Evidence that all hypothalamic hypophysiotropic hormones (HHH) act directly at the level of the pituitary cells is afforded by the demonstration of in vitro experiments on the release or secretion of a given pituitary trophic hormone in culture medium. The addition of LRF to cultured pituitary cells resulted in a high rate of secretion that remained constant for at least a 4-hr period, indicating that LRF is capable of stimulating pituitary cells as long as they are present in the incubation medium (7). The initial pituitary responses to HHH involve the release of stored hormones. Short term pretreatment of the pituitary with puromycin or cyclohexamide does not decrease the amount of hormone secreted in response to HHH (30), indicating that continuous de novo synthesis of the appropriate hormone is not required for the acute release in response to HHH.

There is, however, both in vitro and in vivo evidence that LRF also influences the rate of synthesis of pituitary LH and FSH; addition of nanogram amounts of LRF (once daily for five days) to female rat anterior pituitary cultures significantly increased the release of both LH and FSH. In addition the total content of LH and FSH in pituitary tissue (or cells) and medium exceeded that of controls (31, 32). Similar evidence was obtained in vivo during chronic administration of LRF; in hypophysectomized rats bearing pituitary grafts, chronic treatment with LRF stimulated the gonadotrophs which was more pronounced for FSH than LH in both sexes (33, 34). It is possible that the increased rate of secretion might constitute the signal which mediates an increased rate of synthesis.

There is evidence which indicates that the adenyl cyclase and cyclic AMP system is a mediator of the action of LRF and other hypothalamic hormones (35).

The structure of LRF and its configuration is required for a full biologic expression; no single amino acid in LRF is absolutely required for binding to the receptor, instead many parts of the molecule are necessary for high potency. There are hundreds of LRF analogues that have been synthesized. Deletion of one amino acid from the structure of native LRF has resulted in substantial if not complete loss of biological expression (7). Substitution and modification of amino acids have been shown to increase LRF activity; (D-Ala^6) LRF has a potency of 450% that of LRF (36); des-Gly-amide^{10}-(pro-N-ethylamide^9) possesses 3–5 times more activity than LRF (37), and a prolonged action (38). These analogues are agonists. Analogues with antagonistic action to LRF have been described; des-His^2-LRF with low intrinsic activity does have affinity for the LRF receptor as evidenced by its inhibition of LH release due to LRF, indicating that histidine is required for stimulation of hormone secretion but not for receptor recognition. des-His^2-LRF behaves as a competitive antagonist, with its effect being overcome by increasing doses of LRF (7, 39). Incorporating the D-Ala^6 modification or des-Gly^{10}-pro-ethylamide modification into des-His^2-LRF enhances its antagonist potency by threefold (40, 41), but these antagonists are still rather weak inhibitors; a 700-fold molar excess of des-His^2-(D-Ala^6) LRF is required to inhibit the LH secretion by 50% in vitro (41).

A novel approach to the development of hypothalamic contraceptives by the future discovery of more potent LRF antagonists will obviously be a major advance.

BINDING OF LRF TO PITUITARY RECEPTOR SITES

Using high specific activity (^{3}H-pro^9) LRF, two populations of binding sites have been found in the rat pituitary cells (42); one with high affinity-low capacity and the other with low affinity-high capacity to bind LRF. The latter is less specific but LRH binding at the high affinity receptor site is competitively inhibited by LRF, LRF agonists, and LRF antagonists, even in the presence of an excess of inactive LRF analogues which saturate the low affinity binding site (7, 42). Thus, the limiting factor for the biologic activities of the analogues is the affinity for the LRF receptor.

LRF ACTIONS IN MAN

Since the initial report by Kastin et al (43) that highly purified porcine LRF could stimulate the release of LH and FSH in humans, the recent availability of synthetic LRF and the recent advances in the regulation of gonadotropin secretion have generated an enormous number of clinical investigations of LRF actions in man both in health and disease. Highlights of the current status of the studies and clinical uses of LRF in humans will be reviewed.

SPECIFICITY

Until recently, one of the concepts in neuroendocrinology held that a "unitary" relationship existed between hypothalamic and its respective pituitary hormones. Each specific hypothalamic hormone was believed to control a single pituitary hormone release. During the past three years, exceptions have been found for all three synthetic hypothalamic hormones; TRF releases prolactin as well as TSH (44), LRF releases FSH as well as LH, and somatostatin inhibits the release of TSH as well as growth hormone (45). In addition, the latter hormone has shown to inhibit the α and β cell function of the pancreas. Like the native hypothalamic decapeptide, synthetic LRF stimulates a greater release for LH than for FSH in all species studied. The identification of only one releasing factor led to difficulties in interpreting the control systems underlying the changing levels of the two gonadotropins, especially during the menstrual cycle. Several possibilities have been raised, including that there exists a separate FSH-releasing factor (FRF) which shows a greater intrinsic activity for the release of FSH than LH (46). Until further characterization has appeared, the existence of such an FRF remains open. Another possibility is that LRF is the sole hypothalamic stimulator of both LH and FSH secretion and that neural substances and gonadal steroids have differential effects on their release—a modulatory effect (4). This view is supported by the histoimmunological evidence that LH and FSH are contained in the same cell population of the human adenohypophysis (47).

LRF RESPONSES IN ADULT MALES

In order to avoid cyclic variation, several aspects of pharmacodynamics and physiological responses to synthetic LRF were characterized in normal eugonadal males. Based on the log-dose response relationship studies, 100–150 μg can be used as "near maximal" stimulatory dose and 10 μg is regarded as a threshold dose for repeated challenge in the assessment of pituitary sensitivity (48). There is a high degree of variability in LRF responses among normal subjects (49–51), and it is not related to age or body surface area (49). A relative reproducibility in responses to LRF (10- or 150-μg doses) in subjects with an intact hypothalamic-pituitary-testicular axis can be obtained (52). When responding, an increase in circulating levels of LH occurred within 2 min. The maximal increments (Δ) were greater and reached sooner for LH (25 min) than for FSH (45 min). The decline following the peak response was relatively slow as compared to the rise; the approximate rate of decline had a mean T ½ of 92 min for LH and 450 min for FSH (49). These rates of disappearance in gonadotropin levels following LRF stimulation were 2–3 times longer than those previously estimated following hypophysectomy in humans. This disparity is likely related to the contribution of a continued basal secretion during the disappearance of that additional component of circulating gonadotropin induced by LRF (49, 53). In any event, gonadotropin levels have returned to normal range within 6 hr. There was no change in circulating levels of TSH, prolactin, or growth hormone, indicating the specificity of the gonadotropin releasing activity of synthetic LRF. Because of the episodic fluctuation of circulating testosterone and because of the short lasting effect of LRF on gonadotropin rise, a consistent testicular response as reflected by the increase in circulating testosterone was not found (54, 55). The time course of gonadotropin responses to LRF are depicted in Figure 1. When compared with the low estrogen phase of the menstrual cycle (early follicular phase, $D_{2\text{-}5}$), normal males exhibited a significantly greater maximal re-

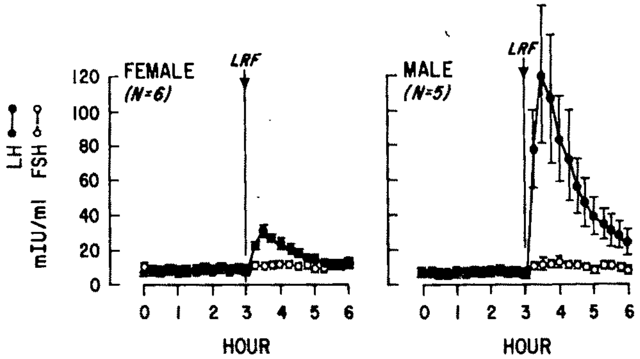


Figure 1 Serum LH and FSH responses to LRF (150 μg) in a normal adult male and female (early follicular phase).

lease of LH but not FSH. These dynamics provide the essential information for the assessment of cyclic variation in normal females, hypogonadal subjects, and patients with hypothalamic-pituitary disorders (see below).

Comparison of different routes of LRF administration have mainly been evaluated in normal male subjects. Following 100 μg of synthetic LRF given intravenously, intramuscularly, or subcutaneously, the release of LH and FSH was quantitatively similar with an identical time course (53, 55). The rise and fall in LH and FSH in circulation lasted approximately 4–6 hr irrespective of the route of administration. A more sustained LH and FSH release can be accomplished by i.v. infusion (55).

Intranasal application of a single dose of synthetic LRF dissolved in saline has shown to be active but the relative potency is much less when compared with other routes of administration (53). In view of the efficacy of a 100-μg dose following subcutaneous injection, a preferable means of self-administration by the patient at six hourly intervals for chronic treatment appears to be a reasonable alternate in the absence of long acting preparation.

LRF RESPONSIVENESS DURING THE MENSTRUAL CYCLE AND EFFECTS OF EXOGENOUS STEROIDS

In contrast to normal adult males, a remarkable cyclic variation of pituitary gonadotropin responsiveness to a given dose of LRF was disclosed (56–58). There was a progressive increase in ΔLH responses to LRF between early and late follicular and mid-luteal phases of the ovulatory cycle. The largest ΔLH was found during mid-luteal phase and the largest ΔFSH was found during late follicular phase. Two additional features were observed; a delayed and more sustained LH and FSH response was revealed during the late follicular phase and a "window" of maximal pituitary sensitivity was found for both LH and FSH during the spontaneous midcycle surge (Figure 2). It is likely that these events were brought about by the direct feedback action of increasing levels of serum estradiol from the mature follicles and the corpus luteum. This assumption is supported by the recent demonstration that the variation of pituitary responsiveness to LRF during the course of ovulatory cycle can be completely negated by the administration of the anti-estrogen clomiphene (Figure 3) (59). Moreover, both negative and positive feedback action of estradiol on pituitary sensitivity to LRF, which are time- and dose-related events, have been observed (60). The quantitative and qualitative differences between responses seen during late follicular phase and the mid-luteal phase may reflect a synergistic action of progesterone and estrogen in further augmentation of pituitary sensitivity to LRF. Using multiple small doses (10 μg every 2 hr X 5) as a stimulus, such a synergistic action has been revealed following progesterone administration (10 mg i.m.) in estrogen-primed normal subjects during the early follicular phase (61). Thus, the low basal gonadotropin levels found during the luteal phase are most likely due to a decreased endogenous LRF secretion occasioned by the synergistic feedback action of estrogen plus progesterone rather than a diminished pituitary sensitivity to LRF.

Although these studies illustrate the positive feedback action of estradiol on pituitary gonadotrophs, the mechanism(s) to account for the acute surge of both LH and FSH at midcycle remain to be resolved. The recent finding that estrogen may concomitantly exert a positive feedback on endogenous LRF release near midcycle (see above), offers preliminary support for the supposition that the event of cyclic gonadotropin release which leads to the acute surge at midcycle is regulated by a dual mechanism of a progressive increase in pituitary sensitivity to LRF and an enhanced LRF secretion.

Due to the brief gonadotropin rise in response to a single injection of LRF, measurable changes in circulating ovarian steroids were not found (Figure 2). On the other hand, repeated LRF injections (twice or more) during the follicular phases of the cycle resulted in a bizarre pattern in the daily fluctuation of gonadotropin and ovarian steroids. While spontaneous midcycle surge and apparent ovulation have occurred, departures from normal are evident; inappropriate surges of LH with or without concomitant rise of FSH during the luteal phase, irregular large fluctuation of estradiol, and an early fall of progesterone but not estradiol several days before the onset of menses were frequently observed (52). An explanation to account for these complex events is not available at the present time. However, based on these

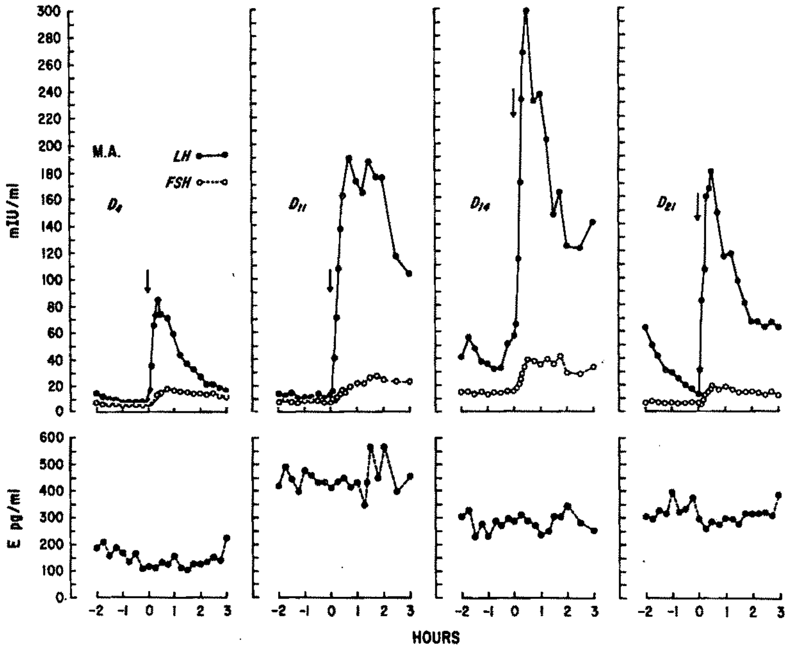


Figure 2 Serum LH, FSH, and estradiol (E) responses to LRF (150 μg) during different phases of the menstrual cycle.

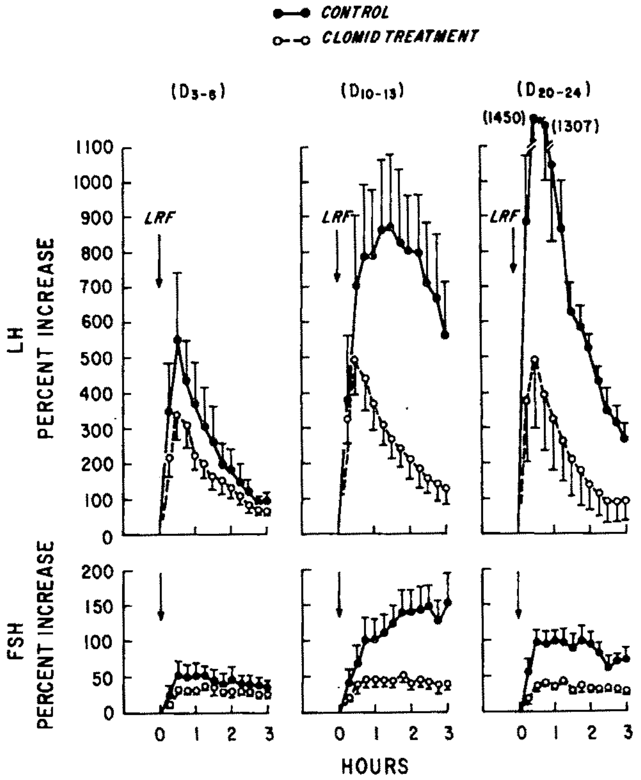


Figure 3 Effect of clomiphene (Clomid ®) on the variation of gonadotropin responses to LRF during the menstrual cycle (data expressed as percent increase from baseline at zero time).

findings, it is reasonable to predict that the contemplated application of LRF-antagonists to block the midcycle surge of LH may similarly disrupt the coordinated sequence of interdependent events between the hypothalamic-pituitary axis and ovarian function.

Assessment of the effects of contraceptive steroid on the pituitary responsiveness to LRF has revealed a marked contrast in the time course and maximal rise of LH in response to the same dose of LRF (150 μg) between subjects receiving sequential and combination cyclic estrogen and progestin treatment (62); a prompt but quantitatively smaller LH release was seen during cyclic combination therapy, while a delayed (5 times) but enhanced (5-fold) LH response was observed during estrogen segments of cyclic sequential therapy. These findings are consistent with the other experimental evidence which indicates the augmentory action of estrogen on pituitary sensitivity to LRF. The failure to observe a synergistic action in subjects receiving combination of estrogen and progestin may be related to the pharmacological nature of the event.

PITUITARY-GONADAL RESPONSIVENESS TO LRF DURING THE PERIOD OF SEXUAL MATURATION

The results of several studies (63–67) indicate that the release of LH following the administration of LRF is minimal in prepubertal children, increasing strikingly at puberty, and a further increase is found in adult males. There is a sex difference in FSH release and in dose-response relationship (64); prepubertal and pubertal females release significantly more FSH than males at all stages of sexual maturation. Hence, there is a striking discordance of the readily releasable LH and FSH between prepubertal and pubertal individuals as well as a sex difference in readily releasable pituitary FSH. In addition, prepubertal females did not exhibit a progressive increase in the release of LH and FSH over a dose range from 1 $\mu g/m^2$ to 100 $\mu g/m^2$ of LRF as was found in prepubertal boys. These findings suggest that the pituitary gonadotrophs of prepubertal females are more sensitive to LRF than those of prepubertal males, even though there is no apparent difference in the concentration of circulating gonadal steroids at this stage of sexual maturation.

The transient increase in LH and FSH in response to LRF was accompanied by a significant increase in plasma testosterone levels in pubertal (but not prepubertal) boys which was significantly less than the incremental testosterone found in adult males (66). These findings are consistent with previous proposals of a sequence of hypothalamic→pituitary→gonadal maturations in human puberty.

The recent disclosure of the sleep-associated increase in episodic pituitary LH secretion which occurs selectively at the time of puberty has advanced our understanding of the pubertal event (68–70). There is a progressive increase in the episodic release of LH during sleep from puberty stage I through stage V. A temporally related testosterone secretion during sleep-induced LH rise is also found (71). Since this augmentation of LH release during sleep is independent of gonadal activity, as it is found in gonadal dysgenesis, a renewed concept of the "CNS program" for the activation of the hypothalamic-pituitary system as a critical step in the initiation of puberty has been proposed (68, 69). This centrally dominated mechanism is further emphasized by the fact that the pituitary responsiveness to LRF in agonadal subjects below age eleven is noticeably smaller than in agonadal subjects after age twelve. The latter exhibited an increased release in the basal state and an augmented response to LRF for LH and FSH (72). Thus, taken together, puberty is initiated by the progressive maturation of the "CNS-entrained program," which manifests itself only during sleep and disappears after puberty. It is likely that this program represents a major, if not the sole input for the augmentation of hypothalamic LRF synthesis and release. More information concerning the biologic rhythm in general and CNS-puberty rhythm in particular are surely forthcoming.

PITUITARY RESPONSIVENESS IN PATIENTS WITH ABNORMAL HYPOTHALAMIC-PITUITARY-GONADAL AXIS

Hypogonadism

Hypogonadism due to a variety of causes such as agonadism, gonadal dysgenesis, germinal aplasia in man, premature ovarian failure, postmenopause, and castration are associated with hypergonadotropinism. The pituitary responsiveness to LRF in these patients is exaggerated (64, 72, 73). This finding indicates that in the absence of gonadal feedback, some degree of physiologic restraint of FSH and LH secretion remains operative and that an increased pituitary store for readily releasable LH and FSH occurs in the absence of gonadal feedback. Similarly, the presence of an inappropriate elevation of gonadotropin in association with germinal aplasia (74), Klinefelter's syndrome (high FSH and relatively normal LH), and polycystic ovarian syndrome (high LH and normal FSH) (61), was accompanied by an exaggerated LRF response for both LH and FSH. These findings provide critical information for future delineation of mechanism(s) in feedback regulation of gonadotropin secretion.

Abnormal Hypothalamic-Pituitary System

The presence of hypogonadotropism raises three immediate questions as to whether the cause is related to (*a*) inability to release gonadotropin by the adenohypophysis, (*b*) inadequate synthesis due to chronic deficiency of the hypothalamic stimulation, or (*c*) inadequate synthesis due to the destruction of pituitary gonadotropin-producing cells. Although the availability of synthetic LRF has afforded a direct functional evaluation of the gonadotroph, data obtained thus far indicate that discriminating hypothalamic vs pituitary origin and functional vs organic lesion is frequently difficult.

ISOLATED GONADOTROPIN DEFICIENCY Gonadotropin responses to LRF in patients with isolated gonadotropin deficiency (IGD) were highly heterogenous (75–77). The characteristically low basal gonadotropin levels in these patients independent of sex or the presence or absence of anosmia responded to LRF in a variable fashion; the majority of patients showed a rise in both FSH and LH. Others showed either an LH response or no change in either hormone. Thus, the syndrome of IGD appears to be heterogenous, although in most patients the pituitary gonadotrophs are intact and the defect most likely resides in the synthesis and secretion of hypothalamic LRF.

SHEEHAN'S SYNDROME The pituitary infarction due to vascular constriction associated with postpartum hemorrhage and shock originally described by Sheehan also showed a heterogenous response to LRF. In the few patients studied, some responded normally, others exhibited either a small or no response (78). This is not unexpected since the degree of infarction and the rate of regeneration may be variable from patient to patient. Thus, LRF testing offers a useful prognostic approach.

AMENORRHEA The absence of menses in an otherwise normal woman represents a most common but difficult diagnostic problem. Clearly, application of LRF testing has provided a powerful tool in assessing the integrity of the hypothalmic-pituitary system (79–81); the majority of amenorrhea patients with either normal or low basal gonadotropin secretion exhibited normal or exaggerated responses for both LH and

FSH, indicating a common etiological factor of hypothalamic dysfunction, the so-called hypothalamic amenorrhea. In patients with anorexia nervosa, a variable response to LRF has been observed; in most patients, the responses were within normal range, indicating an intact pituitary (82).

HYPOPITUITARISM AND PITUITARY TUMOR In both conditions, LRF responses are either reduced or subnormal but in few patients, the gonadotropin responses to LRF have been shown to be within normal range (77, 80, 83). Thus, the diagnostic acuity by LRF responses alone in these conditions is necessarily limited. For this reason, a sequential stimulation by arginine infusion (or insulin) and LRF plus TRF injection for the release of five pituitary hormones (LH, FSH, growth hormone, prolactin, and thyroid stimulating hormone) as an integrated functional reserve test of the adenohypophysis was devised (80, 84, 85); preliminary data indicate that this test has afforded a direct assessment of pituitary reserve and an indirect evaluation of hypothalamic function, particularly in patients with amenorrhea galactorrhea syndrome. It is apparent that in conjunction with other clinical and laboratory data, this test may enhance our ability to distinguish the basic defect of functional vs neoplastic nature (84), hypothalamic vs pituitary origin (85) with single or multiple hormonal involvement.

THERAPEUTIC USES

Because of the short half-life of synthetic LRF, its application in the induction of ovulation and for chronic treatment such as in patients with isolated gonadotropin deficiency is limited. Furthermore, the inherent property of LRF of a greater release of LH than FSH makes the induction of folliculogenesis in most patients with amenorrhea less desirable. The future clinical application of LRF as specific treatment and as a therapeutic adjunct remains to be explored.

Literature Cited

1. Green, J. D., Harris, G. 1947. *J. Endocrinol.* 5:136
2. Matsuo, H., Baba, Y., Nair, R. M. G., Arimura, A., Schally, A. V. 1971. *Biochem. Biophys. Res. Commun.* 43:1374
3. Burgus, R. et al 1972. *Proc. Nat. Acad. Sci. USA* 69:278
4. Blackwell, R. E., Guillemin, R. 1973. *Am. Rev. Physiol.* 35:357
5. Schally, A. V., Kastin, A. J., Arimura, A. 1973. *Vitam. Horm. (New York)* 30:83
6. Schally, A. V., Arimura, A., Kastin, A. J. 1973. *Science* 179:341
7. Vale, W., Grant, G., Guillemin, R. 1973. In *Frontiers in Neuroendocrinology*, ed. W. F. Ganong, L. Martini, 397. New York: Oxford Univ.
8. Arimura, A. et al 1973. *Endocrinology* 93:1092
9. Nett, T. M., Akbar, A. M., Niswender, G. D., Hedlund, M. T., White, W. F. 1973. *J. Clin. Endocrinol. Metab.* 36:880
10. Barry, J., Dubois, M. P., Puolain, P., Leonardelli, J. 1973. *C. R. Acad. Sci.* 276:3191
11. Zimmerman, E. A., Hsu, K. E., Ferin, M., Kozlowski, G. P. 1974. *Endocrinology* 95:1
12. Kumar, T. E. A. 1973. *J. Reprod. Fert. Suppl.* 20:11
13. Kamberi, I. A., Mical, R. S., Porter, J. C. 1971. *Endocrinology* 89:1042
14. Ben-Jonathan, N., Mical, R. S., Porter, J. C. 1973. *Endocrinology* 93:497
15. Seyler, L. E. Jr., Reichlin, S. 1973. *J. Clin. Endocrinol. Metab.* 37:197
16. Malacara, J. M., Seyler, L. E. Jr., Reichlin, S. 1972. *J. Clin. Endocrinol. Metab.* 34:271

17. Arimura, A., Kastin, A. J., Schally, A. V. 1974. *J. Clin. Endocrinol. Metab.* 38:510
18. Seyler, L. E. Jr., Canalis, E., Reichlin, S. 1974. *Clin. Res.* 22:349A
19. Seyler, L. E. Jr., Reichlin, S. 1974. *J. Clin. Endocrinol. Metab.* 39:906
20. Fuxe, K., Hokfelt, T. 1969. In *Frontiers in Neuroendocrinology*, ed. W. F. Ganong, L. Martini New York: Oxford Univ.
21. Bjorklund, A., Falck, B., Nromek, K., Owman, C. H., West, K. A. 1970. *Brain Res.* 17:1
22. Ondo, J. E., Mical, R. S. 1972. *Endocrinology* 31:1239
23. Cuello, A. C., Weiner, R., Ganong, W. F. 1973. *Brain Res.* 59:191
24. Yen, S. S. C., VandenBerg, G., Tsai, C. C., Siler, T. 1974. In *Biorhythms and Reproduction*, ed. M. Ferin et al, Chap. 13, 203. New York: Wiley
25. Seyler, L. E. Jr., Reichlin, S. 1974. *J. Clin. Endocrinol. Metab.* 39:471
26. Redding, T. W. et al 1973. *J. Clin. Endocrinol. Metab.* 37:626
27. Yukitaka, M., Mecklenburg, R. S., Hansen, J. W., Lipsett, M. B. 1973. *J. Clin. Endocrinol. Metab.* 37:63
28. Keye, W. R., Kelch, R. P., Niswender, G. D., Jaffe, R. B. 1973. *J. Clin. Endocrinol. Metab.* 36:1263
29. White, W. F. et al 1974. *Endocrinology* 94:1422
30. Vale, W., Burgus, R., Guillemin, R. 1968. *Neuroendocrinology* 3:34
31. Labrie, F. et al 1973. In *Karolinska Symposium on Research Methods in Reproductive Endocrinology*, ed. E. Diczfalusy, 301. Stockholm: Karolinska Inst.
32. Redding, T. W., Schally, A. V., Arimura, A., Matsuo, H. 1972. *Endocrinology* 90:764
33. Debeljuk, L., Arimura, A., Shiino, M., Rennels, E. G., Schally, A. V. 1973. *Endocrinology* 92:921
34. Arimura, A., Debeljuk, L., Shiino, M., Rennels, E. G., Schally, A. V. 1793. *Endocrinology* 92:1507
35. Borgeat, P. et al 1972. *Proc. Nat. Acad. Sci. USA* 69:2677
36. Monahan, M., Amoss, M., Anderson, H., Vale, W. 1973. *Biochemistry* 12:4616
37. Fujino, M. S. et al 1973. *J. Med. Chem.* 16:1144
38. Rippel, R. H. et al 1973. *Endocrinology* 93:1449
39. Vale, W. et al 1973. *Science* 176:9333
40. Coy, D. J., Vilchez-Martinez, J. A., Coy, E. J., Arimura, A., Schally, A. V. 1973. *J. Clin. Endocrinol. Metab.* 37:331
41. Amoss, M. S. Jr., Monahan, M., Rivier, J., Guillemin, R. 1974. *Endocrinology* 94:A–78
42. Grant, G., Vale, W., Rivier, J. 1973. *Biochem. Biophys. Res. Commun.* 50:771
43. Kastin, A. J., Schally, A. V., Gual, C., Midgley, A. R. Jr., Bowers, C. Y. 1969. *J. Clin. Endocrinol. Metab.* 29:1046
44. Bowers, C. Y., Friesen, H. G., Folkers, K. 1973. *Biochem. Biophys. Res. Commun.* 51:512
45. Yen, S. S. C. et al 1974. *NIH Growth Hormone Symposium, Baltimore, Md.*, ed. S. Raiti, 609. Washington, DC: GPO
46. Bowers, C. Y., Currie, B. L., Johansson, K. N. G., Folkers, K. 1973. *Biochem. Biophys. Res. Commun.* 50:20
47. Phifer, R. F., Midgley, A. R. Jr., Spicer, S. S. 1973. *J. Clin. Endocrinol. Metab.* 36:125
48. Rebar, R. et al 1973. *J. Clin. Endocrinol. Metab.* 36:10
49. Yen, S. S. C. et al 1972. *J. Clin. Endocrinol. Metab.* 34:1108
50. Kastin, A. J., Schally, A. V., Gual, C., Arimura, A. 1972. *J. Clin. Endocrinol. Metab.* 34:753
51. Besser, G. M. et al 1972. *Brit. Med. J.* 3:267
52. Yen, S. S. C. et al 1973. In *Hypothalamic Hypophysiotropic Hormones (Physiological and Clinical Studies)*, ed. C. Gual, E. Rosemberg, 167. Amsterdam: Excerpta Med.
53. Mortimer, C. H., Besser, G. M., Hook, J., McNeilly, A. S. 1974. *Clin. Endocrinol.* 3:19
54. Judd, H., Rebar, R., VandenBerg, G., Yen, S. S. C. 1974. *J. Clin. Endocrinol. Metab.* 38:8
55. Gonzalez-Barcena, D. et al 1973. *J. Clin. Endocrinol. Metab.* 37:481
56. Yen, S. S. C., Vandenberg, G., Rebar, R., Ehara, Y. 1972. *J. Clin. Endocrinol. Metab.* 35:897
57. Nillius, S. J., Wide, L. 1972. *J. Obstet. Gynaecol. Brit. Commonw.* 79:861
58. Thomas, K., Cardon, M., Donnez, J., Ferin, J. 1973. *Contraception* 7:289
59. Wang, C. F., Yen, S. S. C. 1975 *J. Clin. Invest.* 55:201
60. Yen, S. S. C., VandenBerg, G., Siler, T. M. 1974. *J. Clin. Endocrinol. Metab.* 39:170

61. Yen, S. S. C., Lasley, B. L., Wang, C. F., Leblanc, H. 1975. *Recent Progr. Horm. Res.* In press
62. VandenBerg, G., DeVane, G., Yen, S. S. C. 1974. *J. Clin. Invest.* 53:1750
63. Job, J. C., Garnier, P. E., Chaussain, J. L., Milhaud, G. 1972. *J. Clin. Endocrinol. Metab.* 35:473
64. Roth, J. C., Kelch, R. P., Kaplan, S. L., Grumbach, M. M. 1972. *J. Clin. Endocrinol. Metab.* 35:926
65. Thomas, K., Malvaux, P., Ferin, J. 1972. *J. Clin. Endocrinol. Metab.* 35:938
66. Roth, J. C., Grumbach, M. M., Kaplan, S. L. 1973. *J. Clin. Endocrinol. Metab.* 37:680
67. Chaussain, J. L. et al 1974. *J. Clin. Endocrinol. Metab.* 38:58
68. Boyar, R. et al 1972. *N. Engl. J. Med.* 287:582
69. Boyar, R. et al 1973. *J. Clin. Endocrinol. Metab.* 37:521
70. Judd, H. L., Parker, D. C., Siler, T. M., Yen, S. S. C. 1974. *J. Clin. Endocrinol. Metab.* 38:710
71. Parker, D. C., Judd, H., Rossman, L. G., Yen, S. S. C. 1975. *J. Clin. Endocrinol. Metab.* In press
72. Job, J. C. et al 1974. *J. Clin. Endocrinol. Metab.* 38:1109
73. Siler, T. M., Yen, S. S. C. 1973. *J. Clin. Endocrinol. Metab.* 37:491
74. Mecklenburg, R. S., Sherins, R. J. 1974. *J. Clin. Endocrinol. Metab.* 38:1005
75. Bell, J. et al 1973. *J. Clin. Endocrinol. Metab.* 36:791
76. Zarate, A., Soria, J., Canales, E. S., Kastin, A. J., Schally, A. V. 1973. *Obstet. Gynecol.* 42:507
77. Yen, S. S. C., Rebar, R., Vandenberg, G., Ehara, Y., Siler, T. 1973. *J. Reprod. Fert. Suppl.* 20:137–61
78. Aono, T., Minagawa, J., Kinugasa, T., Tanizawa, O., Kurachi, K., 1973. *J. Obstet. Gynecol.* 117:1046
79. Yen, S. S. C., Rebar, R., VandenBerg, G., Judd, H. 1973. *J. Clin. Endocrinol. Metab.* 36:911
80. Harsoulis, P. et al 1973. *Brit. Med. J.* 4:326
81. Nillius, S. J., Wide, L. 1972. *J. Obstet. Gynaecol. Brit. Commonw.* 79:874
82. Mecklenburg, R. S., Loriaux, D. L., Thompson, R. H., Andersen, A. E., Lipsett, M. B. 1974. *Medicine* 53:147
83. Coscia, A. M., Fleischer, N., Besch, P. K., Brown, L. P., Desiderio, D. 1974. *J. Clin. Endocrinol. Metab.* 38:83
84. Rakoff, J., VandenBerg, G., Siler, T. M., Yen, S. S. C. 1974. *Am. J. Obstet. Gynecol.* 119:358
85. Strauss, J. H. et al 1974. *J. Clin. Endocrinol. Metab.* 39:639

Copyright 1975. All rights reserved

HUNTINGTON'S DISEASE

◆7162

Ira Shoulson, M.D. and Thomas N. Chase, M.D.
Laboratory of Neuropharmacology, National Institute of Neurological Diseases and Stroke, National Institutes of Health, Bethesda, Maryland 20014

Huntington's disease (HD) is an autosomal dominant disorder associated with neuronal degeneration especially involving the basal ganglia and cerebral cortex. The cardinal signs of dementia and involuntary movements usually appear in middle life. Although the pathogenesis of this disease remains unknown, its course progressive, and the efficacy of available treatments limited, a recent surge of investigative interest has yielded significant information about HD and related central degenerative disorders.

CLINICAL FEATURES

Most HD patients become symptomatic during the fourth or fifth decades of life with either motor or mental changes or their simultaneous appearance. The diffuse, arrhythmic, semipurposive movements, which constitute the denominative feature of this disorder, characteristically increase during voluntary effort or stress, diminish during repose, and disappear during sleep (1). Associated motor abnormalities may include dysarthria, dysphagia and disturbances in ocular motility (2, 3). Global intellectual impairment constitutes the primary mental feature of HD. Prodromal states include forgetfulness, irritability, restlessness, and personality changes. Affective disorders (euphoria or depression) or psychotic manifestations occur prominently in some patients (4). That thought disorders may occur with HD has suggested an anatomic-biochemical relationship to schizophrenia (5). In addition to the cardinal mental and motor signs of HD, evidence of impaired carbohydrate metabolism (6, 7) and of a defect in the postural vasoregulatory mechanism has recently been reported (8). The mean duration of life from onset of symptoms is about 16 years (9, 10); death usually results from respiratory causes associated with progressive dysphagia (11).

Six to ten percent of HD patients manifest muscular hypertonicity and bradykinesia rather than chorea (9, 12, 13). Although this rigid-hypokinetic (Westphal) variant of HD can develop in adults, it occurs most commonly in patients with onset in the first or second decades of life (12–14). In those with childhood onset, cerebellar ataxia, tremor, dystonic posturing, pyramidal tract signs, and seizures may also occur (15). Mental retardation is often quite severe and may be the earliest clinical sign. Childhood onset HD usually follows a more rapid course than cases with adult onset.

Conventional neurological diagnostic tests such as electroencephalographic or standard cerebrospinal fluid (CSF) studies add little to the evaluation of HD patients. Although the pneumoencephalogram classically reveals dilatation of the anterior horn of the lateral ventricle due to loss of the caudate nucleus bulge (16), the usefulness of this test is limited to the occasional patient with intellectual impairment and motor abnormalities in whom there is no adequate family or past medical history. In view of its rapidly improving quality and availability, the noninvasive, computerized axial tomographic technique may eventually supplant the pneumoencephalogram.

Among the conditions to be considered in the diagnostic evaluation of patients presenting in middle life with bilateral chorea, systemic lupus erythematosus (17), thyrotoxicosis (18), polycythemia vera (19), and hypoparathyroidism (20) should be included. Choreic movements may occur following use of oral contraceptives (21), psychotropic phenothiazine and butyrophenone derivatives, and dopamine receptor agonists such as L-dopa, piribedil, and apomorphine (22). Other neurologic disorders with chorea include paroxysmal choreoathetosis (23) and chronic psychotic choreoathetosis (24). Choreiform movements and late onset dementia have also been reported with acanthocytosis (25, 26).

EPIDEMIOLOGY AND GENETICS

HD has been described in virtually all ethnic and racial groups with a prevalence rate of about 5 per 100,000 (9, 10, 27, 28). Assuming full penetrance of the HD gene, a mutation rate of only 1 in 200,000 chromosomes per generation has been estimated (29). Such a calculation is partially dependent on the fertility or genetic fitness of HD patients. It is surprising that the relative fertility of affected individuals significantly exceeds that of their normal sibs (30, 31). Despite this, HD families produce no more children than the general population because the fertility of unaffected sibs is lower than that of the general population. The foregoing observations might indicate that the deleterious gene is fitter than its normal allele (31). On the other hand, selection of large kindreds and voluntary limitations of reproduction by normal sibs might provide an explanation for the apparent increased fitness.

In adult onset HD the sex ratio is 1:1 as would be predicted on the basis of the autosomal dominant mode of inheritance. On the other hand, paternal descent is more common in cases which begin in childhood or adolescence (32); approximately 75% of *children* with HD inherit the disease from their father. Furthermore, the ratio of female to male childhood patients is 2:1, whereas it is the expected 1:1 in the adult onset cases. A provocative explanation for this phenomenon postulates polygenic modifiers controlling age of onset of HD. The question thus arises whether there are important polyallelic genetic modifiers, perhaps sex linked, accounting for the childhood variant of HD and its predominance of paternal descent.

EARLY DETECTION

In autosomal dominant disorders, unaffected individuals do not risk transmitting the disease to their progeny. Genetic counseling in most such disorders is simplified by the ability to identify affected individuals prior to their reproductive years. In HD, the late onset of symptoms makes the problem much more complex. Ideally, one would like to develop an in utero means of detection so that antenatal intervention might be appropriately considered. Short of this, a method of detection which is applicable prior to the reproductive years would enable suitable genetic counseling to be instituted (33). On the assumption that subtle clinical changes can be detected in the heterozygote carrier prior to the appearance of overt manifestations, psychological and language tests (34), spatial orientation assessment combined with electromyography (35), and recordings of subtle abnormal movements, particularly of the eyes (36), have been performed. The predictive value of these approaches to the early diagnosis of HD remains uncertain.

The application of biochemical techniques to the presymptomatic detection of HD has received some advocacy following the successful use of L-dopa in the treatment of parkinsonism. In sufficient doses, L-dopa induces choreiform dyskinesias in parkinsonian patients. Based on the assumption that patients who will eventually develop HD are more sensitive to L-dopa-induced dyskinesias than individuals who will not develop the disorder, L-dopa has been given to clinically normal individuals at risk for HD. In two studies, somewhat over one third of at-risk subjects developed transient involuntary movements during L-dopa administration (37, 38). Even if long term follow-up studies show that individuals responding to L-dopa in this manner carry an increased risk of having HD, the psychological impact and the ethical issues involved with this test leave much to be desired (39). A recent report that the life span of fibroblasts cultured from the skin of patients with HD is significantly shorter than that from control subjects (40) has raised the possibility of another approach to the early detection of HD.

Efforts are currently being intensified to establish genetic linkage with a relatively common marker. If linkage were to be demonstrated, offspring could be assigned a modified risk of possessing the HD gene dependent on the presence or absence of the marker trait (41). Then presymptomatic, even prenatal, detection would be facilitated. This prospect has recently been accomplished in the linkage confirmation of loci for myotonic dystrophy and ABH secretion (42). Blood Group A has been reported to be significantly associated with HD (43), but this observation awaits confirmation in larger population studies. Although chromosomal number and gross morphology are normal in HD patients (9), new developments in karyotyping and staining techniques might enhance our ability to detect more subtle chromosomal variations (44).

PATHOLOGY

In adult onset HD, the brain is invariably reduced in size and weight due to widespread neuronal loss, affecting especially small neurons in the neostriatum and in the third, fifth, and sixth layers of cerebral cortex. Unlike parkinsonism, the zona compacta of the substantia nigra appears normal, although dendritic loss and thinning of striatal bundles has been noted in the pars reticulata (45). Neuropathologic features of the juvenile and rigid-hypokinetic variants of HD generally resemble those of adult onset cases, although severe gliosis of the globus pallidus with minimal neuronal depletion (46) and more marked neostriatal, cortical, and cerebellar changes have been reported (12, 47).

Ultrastructural studies of cerebral biopsy material from HD patients (48) have shown unusually large lipofuscin accumulation in neurons, increased amounts of smooth endoplasmic reticulum, disrupted mitochondria, abnormal presynaptic endings, and axonal degeneration and proliferation. Adjunctive histochemical studies disclosed large cytoplasmic lipofuscin granules in astrocytes having a strikingly elevated acid phosphatase activity. These astrocytic changes, the accumulation of lipopigments, and the high activity of hydrolytic enzymes suggest that the gliosis characteristically attending the neuronal degeneration in HD may not be entirely "reactive," and that astrocytes as well as neurons may be involved in a more generalized metabolic disorder possibly affecting membranes.

Although pathological changes outside of the central nervous system have not heretofore been thought to accompany HD, a recent study of liver biopsy specimens revealed histologic abnormalities of hepatocytes (49). It was suggested that, like neurons, hepatocytes have a shortened survival capacity. Unlike neuronal tissue, however, hepatocytes have the ability to regenerate and thus to better maintain their functional as well as structural integrity. The foregoing observations suggest that the examination of extraneuronal tissues should not be ignored in future studies of HD pathophysiology.

BIOCHEMISTRY AND PHARMACOLOGY

Dopamine metabolism has recently been accorded considerable investigative attention in HD patients due to the now well documented relationship between neural systems containing this amine and extrapyramidal function. Although a somewhat imperfect comparison, HD has been considered a mirror image of Parkinson's disease. While in parkinsonian patients, L-dopa ameliorates motor signs and produces dose-related dyskinesias, in patients with HD, L-dopa tends to exacerbate involuntary movements. Conversely, drugs which block dopaminergic function either by inhibiting dopamine synthesis (α-methylparatyrosine), intraneuronal storage (reserpine or tetrabenazine), or postsynaptic interactions (phenothiazines or butyrophenone derivatives) are modestly effective in suppressing the motor symtoms of HD (50).

Although steady state levels of homovanillic acid, the principle dopamine metabolite, are variably decreased in the lumbar CSF of patients with HD (51, 52), dopamine turnover, as estimated by the oral probenecid loading technique, shows a consistent and substantial reduction (50). Studies of postmortem brain material from patients with HD also have disclosed a significant reduction in both dopamine and homovanillic acid concentrations in the caudate nucleus (53). The changes in central dopamine metabolism can probably not be attributed to structural alterations in the nigrostriatal dopaminergic pathway, since in contrast to the situation in Parkinson's disease, this system appears to remain morphologically intact in patients with HD. Conceivably, neurons comprising this pathway make synpatic contact with the striatal cells which degenerate in HD. If dopamine acts as an inhibitory transmitter at these synapses, then it is possible that a reduction in dopamine metabolism (and thus presumably in dopamine-mediated transmission) may reflect a compensatory, functional response to the striatal cell degeneration.

Some of the most provocative findings in this disorder have derived from studies of the brain amino acid composition of HD patients. Significantly decreased levels of γ-aminobutyric acid (GABA) and homocarnosine have been reported in the substantia nigra, caudate, putamen, and globus pallidus (54). GABA probably functions as an inhibitory transmitter throughout much of the mammalian central nervous system. A substantial reduction in the activity of glutamic acid decarboxylase, the enzyme mediating GABA synthesis, has also been observed in the brains of HD patients (55, 56). Recent evidence of more diverse intracellular enzyme deficiencies, as well as glutamic acid decarboxylase reductions in parkinsonian patients (57), however, raises some question about the specificity of the decarboxylase findings to HD, and suggests that a more generalized disturbance of cellular enzymes may exist in HD (58).

Alterations in the function of several other neurotransmitter systems have been suggested to occur in HD patients. Reduced CSF concentrations of choline have been reported (59), and postmortem brain specimens from HD patients show a patchy reduction of choline acetylase in the neostriatum despite normal acetylcholinesterase levels (56). This may indicate a relative deficiency state in HD as contrasted with the choline hyperactivity seen in Parkinson's disease. Indeed, improvement in choreiform movements is reported by some investigators following intravenous physostigmine, a centrally active cholinergic agent (60, 61). At present, neither biochemical nor pharmacologic evidence lends very strong support to the contention that serotonin containing neural systems play a critical role in the pathophysiology of HD (62–65).

MANAGEMENT

Among currently available drugs, those which deplete central dopamine stores or block dopamine receptors afford the greatest symptomatic relief. The rauwolfia alkaloids represent the prototype of pharmacologic agents which deplete dopamine and other central monoamines (12, 66). Tetrabenazine, which resembles reserpine in its action, but which is less likely to cause hypotension, also diminishes chorea (67, 68). Neuroleptics such as haloperidol and perphenazine (69) appear to owe their therapeutic efficacy in HD patients to their ability to block dopamine receptors. In addition to the modest improvement in chorea, patients may also benefit from the antipsychotic action of these drugs. No definite therapeutic superiority has been demonstrated between the commonly available neuroleptics. Both chlordiazepoxide

(70) and diazepam (71) also tend to improve motor symptoms, although not to the same extent as the neuroleptics.

The psychiatric management of HD patients is difficult and often unrewarding despite its critical importance to both patient and his family (72). Depression is frequently of a reactive type and can be managed by appropriate supportive therapy. On the other hand, the presence of severe depression or of psychotic behavior may necessitate more intensive psychiatric intervention. Some patients have responded to antidepressant agents; in this regard, the tricyclic antidepressants are preferable since monoamine oxidase inhibitors may cause worsening of dyskinesias (73) due to inhibition of catecholamine degradation.

Genetic counseling has recently assumed a greater scope with the advent of antenatal diagnosis. Unfortunately, antenatal diagnosis or even the early detection of gene carriers has yet to be realized in HD patients, or for that matter in most Mendelian dominant disorders. With this in mind and with the variety of reproductive controls now available, the dilemma of the genetic counselor becomes more complex. Adequate eugenic guidelines are difficult to establish and it may be argued that genetic counseling of HD patients should be limited to providing factual and useful information regarding all aspects of the disease. Indeed, as a minimum obligation, all family members at risk who make inquiry should be provided with information regarding the clinical course, prognosis, and mode of inheritance of HD. Furthermore, the probability of manifest disease happening to the counselee or other members of his family should be clearly presented. Controversy then arises as to the degree of intervention assumed by the physician or counselor regarding reproduction. Until practical advances are made in the early detection of HD carriers, this critical area will require resolution within the vague convergence of the eugenist and the at-risk patient.

Literature Cited

1. Duvoisin, R. 1972. Clinical diagnosis of the dyskinesias. *Med. Clin. N. Am.* 56:1321–41
2. Dix, M. R. 1970. Clinical observations upon the vestibular responses in certain disorders of the central nervous system. *Advan. Otorhinolaryngol.* 17:118–28
3. Starr, A. 1967. A disorder of rapid eye movements in Huntington's chorea. *Brain* 90:545–64
4. Garron, D. C. 1973. Huntington's chorea and schizophrenia. *Advances in Neurology, Vol. 1: Huntington's Chorea, 1872–1972,* ed. A. Barbeau, T. Chase, G. Paulson. 729–34. New York: Raven
5. Klawans, H. L., Goetz, C., Westheimer, R. 1972. Pathophysiology of schizophrenia and the striatum. *Dis. Nerv. Syst.* 33:711–19
6. Podolsky, S., Leopold, N. A., Sax, D. S. 1972. Increased frequency of diabetes mellitus in patients with Huntington's chorea. *Lancet* 1:1356–58
7. Podolsky, S., Leopold, N. A. 1974. Growth hormone abnormalities in Huntington's chorea: Effect of L-dopa administration. *J. Clin. Endocrinol. Metab.* 39:36–39
8. Aminoff, M. J., Gross, M. 1974. Vasoregulatory activity in patients with Huntington's chorea. *J. Neurol. Sci.* 21:33–38
9. Myrianthopoulos, N. C. 1966. Huntington's chorea. *J. Med. Genet.* 3:298–314
10. Reed, T. E., Chandler, J. H. 1958. Huntington's chorea in Michigan. 1. Demography and genetics. *Am. J. Hum. Genet.* 10:201–25
11. Edmonds, C. 1966. Huntington's chorea, dysphagia and death. *Med. J. Aust.* 63:273–74
12. Bruyn, G. W. 1968. Huntington's chorea: Historical, clinical and laboratory

synopsis. *Handbook of Clinical Neurology,* 6:298–378. Amsterdam: North-Holland

13. Dewhurst, K., Oliver, J. 1970. Huntington's disease of young people. *Eur. Neurol.* 3:278–89
14. Markham, C. H., Knox, J. W. 1965. Observations on Huntington's chorea in childhood. *J. Pediat.* 67:46–57
15. Green, J. B., Dickinson, E. S., Gunderman, J. R. 1973. Epilepsy in Huntington's chorea. See Ref. 4, pp. 105–13
16. Blinderman, E. E., Weidner, W., Markham, C. H. 1964. The pneumoencephalogram in Huntington's chorea. *Neurology* 14:601–7
17. Greenhouse, A. H. 1966. On chorea, lupus erythematosus, and cerebral arteritis. *Arch. Intern. Med.* 117:389–93
18. Fidler, S. M., O'Rourke, R. A., Buchsbaum, H. W. 1971. Choreoathetosis as a manifestation of thyrotoxicosis. *Neurology* 21:55–57
19. Ashenhurst, E. M. 1972. Chorea complicating polycythemia rubra vera. *Can. Med. Assoc. J.* 107:434–37
20. Hossain, M. 1970. Neurological and psychiatric manifestations in idiopathic hypoparathyroidism. *J. Neurol. Neurosurg. Psychiat.* 33:153–55
21. Gamboa, E. T., Isaacs, G., Harter, D. H. 1971. Chorea associated with oral contraceptive therapy. *Arch. Neurol.* 25:112–14
22. Chase, T. N. 1972. Drug-induced extrapyramidal disorders. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 50:448–71
23. Jung, S., Chen, K., Brody, J. A. 1973. Paroxysmal choreoathetosis. *Neurology* 23:749–55
24. Dincmen, K. 1966. Chronic psychotic choreo-athetosis. *Dis. Nerv. Syst.* 27:399–402
25. Critchley, E. M., Clark, D. B., Wikler, A. 1968. Acanthocytosis and neurological disorder without beta lipoproteinemia. *Arch. Neurol.* 18:134–40
26. Cedarbaum, S., Heywood, D., Aigner, R., Motulsky, A. 1971. Progressive chorea, dementia and acanthocytosis. *Clin. Res.* 19:177 (Abstr.)
27. Myrianthopoulos, N. C. 1973. Huntington's chorea: The genetic problem five years later. See Ref. 4, pp. 149–59
28. Stevens, D. L. 1973. Heterozygote frequency for Huntington's chorea. See Ref. 4, pp. 191–98
29. Stevens, D., Parsonage, M. 1969. Mutation in Huntington's chorea. *J. Neurol. Neurosurg. Psychiat.* 32:140–43
30. Reed, T. E., Neel, J. V. 1959. Huntington's chorea in Michigan. 2. Selection and Mutation. *Am. J. Hum. Genet.* 11:107–36
31. Wallace, D. C., Parker, N. 1973. Huntington's chorea in Queensland. See Ref. 4, pp. 223–35
32. Bird, E. D., Caro, A. J., Pilling, J. B. 1974. A sex related factor in the inheritance of Hungtington's chorea. *Ann. Hum. Genet.* 37:255–60
33. Goodman, R. M., Ashkenazi, Y. E., Adam, A., Greenfield, G. 1973. Thoughts on the early detection of Huntington's chorea. See Ref. 4, pp. 273–85
34. Palm, J. D. 1973. Longitudinal study of a preclinical test program for Huntington's chorea. See Ref. 4, pp. 311–24
35. Baro, F. 1973. A neuropsychological approach to early detection of Huntington's chorea. See Ref. 4, pp. 329–38
36. Petit, H., Milbled, G. 1973. Anomalies of conjugate ocular movements in Huntington's chorea. See Ref. 4, pp. 287–94
37. Klawans, H. L., Paulson, G. W., Ringel, S. P., Barbeau, A. 1972. Use of L-dopa in the detection of presymptomatic Huntington's chorea. *N. Engl. J. Med.* 286:1332–34
38. Cawein, M., Turney, F. 1971. Test for incipient Huntington's chorea. *N. Eng. J. Med.* 284:504
39. Lappe, M., Gustafson, J. M., Roblin, R. 1972. Ethical and social issues in screening for genetic disease. *N. Engl. J. Med.* 286:1129–32
40. Menkes, J. H., Stein, N. 1973. Fibroblast cultures in Huntington's disease. *N. Engl. J. Med.* 288:856–57
41. Lindstrom, J. A. et al 1973. Genetic linkage in Huntington's chorea. See Ref. 4, pp. 203–8
42. Renwick, J. H., Bundey, S. E., Ferguson-Smith, M. A. Izatt, M. M. 1971. Confirmation of linkage of the loci for myotonic dystrophy and ABH secretion. *J. Med. Genet.* 8:407–16
43. Schimke, R. N., Ziegler, D. K., 1970. ABO blood groups and Huntington's chorea. *Lancet* 2:475–76
44. Hecht, F., Wyandt, H. E., Erbe, R. W. 1971. Revolutionary cytogenetics. *N. Engl. J. Med.* 285:1482–84
45. Gebbink, T. B. 1968. Huntington's chorea: Fibre changes in the basal ganglia. See Ref. 12, pp. 399–408
46. Byers, R. K., Gilles, F. H., Fung, C., 1973. Huntington's disease in children. *Neurology* 23:561–69
47. Liss, L., Paulson, G. W., Sommer, A.

1973. Rigid form of Huntington's chorea. See Ref. 4, pp. 405–24

48. Tellez-Nagel, I., Johnson, A. B., Terry, R. D. 1973. Ultrastructural and histochemical study of cerebral biopsies in Huntington's chorea. See Ref. 4, pp. 387–98
49. Bolt, J. M., Lewis, G. P. 1973. Huntington's chorea: A study of liver function and histology. *Quart. J. Med.* 42:151–74
50. Chase, T. N. 1973. Biochemical and pharmacologic studies of monoamines in Huntington's chorea. See Ref. 4, pp. 533–42
51. Klawans, H. L. 1971. Cerebrospinal fluid homovanillic acid in Huntington's chorea. *J. Neurol. Sci.* 13:277–79
52. Curzon, G., Gumpert, J., Sharpe, D. 1972. Amine metabolites in the cerebrospinal fluid in Huntington's chorea. *J. Neurol. Neurosurg. Psychiat.* 35: 514–19
53. Bernheimer, H., Birkmayer, W., Hornykiewicz, O., Jellinger, K., Seitelberger, F. 1973. Brain dopamine and the syndromes of Parkinson and Huntington. *J. Neurol. Sci.* 20:415–55
54. Perry, T. L., Hansen, S., Kloster, M. 1973. Huntington's chorea: Deficiency of γ-aminobutyric acid in brain. *N. Engl. J. Med.* 288:337–42
55. Bird, E. D., Iversen, L. L. 1974. Huntington's chorea: Post-mortem measurement of glutamic acid decarboxylase, choline acetyltransferase and dopamine in basal ganglia. *Brain* 97:457–72
56. McGeer, P. L., McGeer, E. G., Fibiger, H. C. 1973. Choline acetylase and glutamic acid decarboxylase in Huntington's chorea. *Neurology* 23:912–17
57. McGeer, P. L., McGeer, E. G., Fibiger, H. C. 1973. Glutamic-acid decarboxylase and choline acetylase in Huntington's chorea and Parkinson's disease. *Lancet* 2:623–24
58. Stahl, W. L., Swanson, P. D. 1974. Biochemical abnormalities in Huntington's chorea brains. *Neurology* 24: 813–19
59. Aquilonius, S. M., Nyström, B., Schuberth, J., Sundwall, A. 1972. Cerebrospinal fluid choline in extrapyramidal disorders. *J. Neurol. Neurosurg. Psychiat.* 35:720–25
60. Klawans, H. L., Rubovits, R. 1972. Central cholinergic-anticholinergic antagonism in Huntington's chorea. *Neurology* 22:107–16
61. Tarsy, D., Leopold, N., Sax, D. S. 1974. Physostigmine in choreiform movement disorders. *Neurology* 24:28–33
62. Tourian, A. Y. 1972. Administration of 5-hydroxy-L-tryptophan to individuals with Huntington's chorea. *Neurology* 22:1201–4
63. Chase, T. N., Watanabe, A. M., Brodie, H. K., Donnelly, E. F. 1972. Huntington's chorea: Effect of serotonin depletion. *Arch. Neurol.* 26:282–84
64. Chase, T. N., Shoulson, I. 1975. Behavioral and biochemical effects of fenfluramine in patients with neurologic disease. *Postgrad. Med. J.* 51(Suppl. 1): 101–5
65. Klawans, H. L., Rubovits, R., Ringel, S. P., Weiner, W. J. 1972. Observations on the use of methysergide in Huntington's chorea. *Neurology* 22:929–33
66. Kempinsky, W. H., Boniface, W. R., Morgan, P. P., Busch, A. K. 1960. Reserpine in Huntington's chorea. *Neurology* 10:38–42
67. Swash, M., Roberts, A. H., Zakko, H., Haethfield, K. W. 1972. Treatment of involuntary movement disorders with tetrabenazine. *J. Neurol. Neurosurg. Psychiat.* 35:186–91
68. McLellan, D. L., Chalmers, R. J., Johnson, R. H. 1974. A double-blind trial of tetrabenazine, thiopropazate, and placebo in patients with chorea. *Lancet* 1:104–7
69. Fahn, S. 1973. Treatment of choreic movements with perphenazine. See Ref. 4, pp. 755–64
70. Warick, L. H., Barrows, H. S. 1964. Treatment of Huntington's chorea with chlordiazepoxide. *Bull. Los Angeles Neurol. Soc.* 29:17–21
71. Farrell, D. F., Hofmann, W. W. 1968. A quantitative evaluation of the effect of diazepam in Huntington's chorea. *Arch. Phys. Med. Rehabil.* 49:586–91
72. Whittier, J. R., Heimler, A., Korenyi, C. 1972. The psychiatrist and Huntington's disease. *Am. J. Psychiat.* 128: 1546–50
73. Whittier, J., Haydu, G., Crawford, J. 1961. Effect of imipramine on depression and hyperkinesia in Huntington's disease. *Am. J. Psychiat.* 118:79

Copyright 1975. All rights reserved

RELATION BETWEEN GROWTH HORMONE AND SOMATOMEDIN[1]

❖7163

Judson J. Van Wyk, M.D.[2] *and Louis E. Underwood, M.D.*[3]
Department of Pediatrics, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27514

DEFINITION OF SOMATOMEDIN

In 1957, Salmon & Daughaday showed that growth hormone stimulates the incorporation of radiolabeled sulfate into proteoglycans of cartilage by inducing a secondary "sulfation factor" (1). Subsequent studies have indicated that normal plasma contains an unknown number of growth hormone dependent substances which stimulate cellular events concerned with the growth of both extraskeletal and skeletal tissues. Although immunologically distinct from insulin, these substances display many insulin-like metabolic actions in vitro. As a result of these and other studies, the original designation "sulfation factor" became entirely too constrictive. The more open-ended term, "somatomedin," was therefore introduced to encompass all growth hormone dependent substances in plasma which stimulate growth in responsive tissues (2).

Although whole plasma stimulates the growth of many tissues in vitro, very little is known concerning the number of specific factors involved, their chemical natures, or their hormonal dependencies. Because of rapid acquisition of new knowledge concerning the somatomedins, it seems wisest at present to avoid overly strict definitions of what substances should or should not be included under this designation. Thus, from our present perspective, it would seem irrational, simply on historical grounds, to exclude from this family of compounds growth hormone dependent peptides which might be inactive in cartilage assays but which nevertheless stimulate

---

[1] Supported by USPHS Research Grants AM01022 and HDO8299, Training Grant AMO5330, and by grants from the Human Growth Foundation, Elsa U. Pardee Foundation, and University of North Carolina Research Council.

[2] JJVW is a recipient of Research Career Award #4 K06 AM14115 from the National Institutes of Health.

[3] LEU is a Jefferson-Pilot Fellow in Academic Medicine, University of North Carolina.

the growth of extraskeletal tissues. Likewise, it now seems wise to drop an earlier suggestion that all members of this family must exhibit insulin-like properties (3, 4). Recent studies now suggest that the growth-promoting effects of somatomedin may be mediated by a different portion of the molecule than are the insulin-like effects. Thus, somatomedins and insulin seem to be related through some ancestral homology in structure which gives rise to a partial overlap of biological action; the growth-promoting actions of the somatomedins are therefore not necessarily dependent on the insulin-like effects.

The purpose of this review is to survey the assay methods which have been used to measure somatomedin levels in plasma and to monitor purification; to review present knowledge of the purification and chemical nature of these substances; to describe their in vitro biological effects; and to review currently available information on the relationship between linear growth, plasma growth hormone, and plasma somatomedin levels. Since most of our present knowledge of the somatomedins has been garnered from studies which have employed some modification of the original sulfation factor assay, the term "somatomedin" will be used in this review to refer to those substances which are detected by these assays. Where the term is used generically, its broader meaning should be obvious from the context.

THE ORIGINAL SOMATOMEDIN PHENOMENON

Salmon & Daughaday (1) reported in 1957 that the incorporation of radioactive sulfate into proteoglycans of cartilage was stimulated when normal serum was incubated with rat costal cartilage and $^{35}SO_4$. When serum from a hypopituitary individual or a hypophysectomized animal was substituted for the normal serum, sulfate incorporation was diminished. Enrichment of the hypopituitary serum by the addition of large quantities of growth hormone caused no change, but the capacity of serum from hypopituitary patients or animals to stimulate sulfate incorporation returned after parenteral administration of growth hormone. From these fundamental observations arose the concept that growth hormone does not stimulate skeletal growth directly, but rather acts by stimulating the formation of a secondary growth-promoting substance. It was subsequently shown that this "sulfation factor" exerted multiple biologic effects on cartilage including the incorporation of thymidine into DNA, uridine into RNA, amino acids into protein, and the conversion of proline to hydroxyproline (5–7). Because of the possibility that different factors might stimulate the synthesis of cartilage matrix and cell replication, several workers have used the operational term "thymidine factor" when measuring the incorporation of thymidine into DNA (8, 9).

ASSAYS FOR SOMATOMEDIN

Cartilage Bioassays

Most bioassay methods for somatomedin are based on the incorporation of $^{35}SO_4$ into cartilage. Cartilage possesses unique advantages for in vitro studies since the predominant cell type is relatively resistant to low oxygen tension and nearly all of the labeled sulfate which is taken up can be identified as sulfated glucosaminoglycans. After a lag phase of 18–20 hr, the incorporation of ^{3}H-methyl thymidine into DNA parallels the incorporation of sulfate (10, 11). Using a dual label assay, Van Wyk has concluded that the plasma factor which stimulates sulfate incorporation is the same as that responsible for the synthesis of DNA (11, 12).

The original sulfation factor assay utilized costal cartilage segments from hypophysectomized rats. This assay has great sensitivity but is exceedingly cumbersome and costly to perform and has large variances. Yde found satisfactory dose response curves in costal cartilages from fasted, immature rats, but this method sacrifices sensitivity due to high basal uptakes of labeled sulfate (13).

The disadvantages of the rat cartilage assay led Hall to introduce a simpler and less time-consuming bioassay which quantitated the incorporation of $^{35}SO_4$ into pelvic leaflets of 11-day chick embryos (14). Although slightly less sensitive than rat cartilage assays, numerous laboratories have reported satisfactory precision with this technique or one of its variations.

Van den Brande et al have recently developed an assay of very high precision using identical pieces of costal cartilage from a single pig (15). These results have been confirmed by Phillips et al, who stressed the importance of using pigs less than 9 months of age (16). The porcine assay is the least sensitive of the cartilage assay systems, but also seems to be the least influenced by inhibitors in plasma.

It must be stressed that all somatomedin bioassays are greatly influenced by inhibitory substances and therefore reflect only net somatomedin activity. Salmon has investigated the nature of these inhibitory substances in rat plasma and found that at least part of the inhibitory material was heat labile, nondialyzable, and destroyed by tryptic digestion (17, 18). This inhibitory material is increased in the serum of fasted and hypophysectomized rats, in hypopituitary patients, and in the plasma of patients with protein-calorie malnutrition (19).

In addition to the variable quantity of macromolecular inhibitors present in sera, the in vitro uptake of $^{35}SO_4$ may be inhibited by fatty acids (20) and high cortisol levels (6, 21, 22). Likewise, failure to correct for variations in the plasma content of inorganic sulfate may, under certain circumstances, give rise to spurious conclusions based on the incorporation of $^{35}SO_4$.

Assay of Somatomedin by Cell Culture Techniques

Most mammalian cells fail to divide in vitro unless their culture medium is enriched with serum or some similar complex of growth factors (23, 24). Gey & Thalhimer (25) demonstrated that insulin at very high dosage would at least partially replace the requirement for serum of cells grown in vitro.

Salmon & Hosse (26) first showed that bovine plasma extracts rich in somatomedin also stimulated the growth of HeLa cells. Uthne has shown that extracts of human plasma which are rich in somatomedin-like activity also stimulate thymidine incorporation by human glial-like cells in culture. With further purification, two separate substances were identified; one stimulated the growth of glial-like cells, whereas the other stimulated sulfate incorporation by embryonic chick pelvic leaflets (27). Both factors qualified as somatomedins, since plasma from hypophysectomized rats was deficient in both activities, and both activities increased following the administration of growth hormone. To differentiate between these substances, he designated the cartilage factor somatomedin-A and the factor which stimulates glial-like cells somatomedin-B.

Somatomedin-rich extracts of human plasma prepared in our laboratory have likewise been found to stimulate thymidine incorporation in human fibroblast cultures, fetal rat liver cell cultures, and in a rat ovarian tumor cell line.[4] Since these effects were obtained with only partially purified preparations of somatomedin, it was not clear which component of this material might be responsible for the mitogenic activity. More recent studies by Dr. Joel Baseman have demonstrated that an essentially pure preparation of somatomedin-C isolated in our laboratory (see below) stimulates dramatic increases in thymidine incorporation and mitotic index in primary cultures of chick fibroblasts. These effects were far greater than the maximal effect which could be achieved with insulin (J. Baseman, unpublished observations).

There are as many variations of cell culture assay techniques as there are laboratories performing them. In general, monolayer cultures of various cell lines are established and grown to confluency. They are then deprived of serum, and the response of test substances is measured in terms of ^{3}H-thymidine incorporation into DNA, increase in mitotic index, or the incorporation of $^{35}SO_4$ into proteoglycans (28). Although the use of cell culture assays to define specific growth factors will undoubtedly be of critical importance in future somatomedin research, it is beyond the scope of this review to summarize the diverse literature on this subject. Much of the research on growth factors has originated in laboratories of cell biology and has been carried out without specific consideration to possible hormonal dependencies.

Specific Radioreceptor Assays for Somatomedin

Specific radioimmunoassays and cell receptor assays will unquestionably replace bioassay techniques when sufficient quantities of the pure somatomedins become available. Marshall et al have developed a radioreceptor assay for somatomedin-C based on the competition between ^{125}I-somatomedin-C and unlabeled somatomedin for binding to cell-free membrane preparations prepared from human placentas (29, 30). The agreement between results obtained with this assay and the rat cartilage bioassay has been excellent. In addition to its immense value in monitoring the purification of somatomedin from partially purified plasma fractions, the placental receptor assay has been found capable of accurately measuring somatomedin levels in unextracted human plasma. Satisfactory dose response curves are obtained at dosages between 1 and 20 μl of plasma per milliliter of incubation medium (Figure 1). This assay has provided good discrimination between the plasma somatomedin levels of hypopituitary, normal, and acromegalic subjects (Figure 2). It also appears

---

[4] Studies of the effects of somatomedin on cultured human fibroblasts were carried out by Michael Swift, M.D., at the University of North Carolina. The fetal liver cell and the rat ovarian tumor cell studies were done by Hyam Leffert, M.D., and Denis Gospodarowicz, Ph.D., respectively, at the Salk Institute, La Jolla, California.

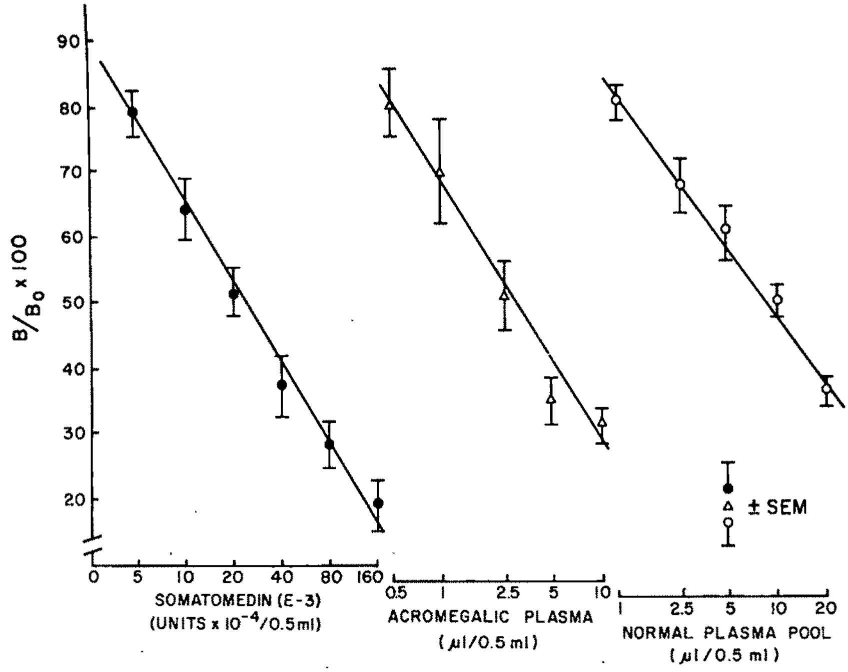


Figure 1 Competition of partially purified somatomedin, acromegalic plasma, and a normal plasma pool for the human placental somatomedin-C receptor. Approximately 150 µg of placental cell membrane protein and tracer quantities of highly purified, radioiodinated somatomedin-C were incubated overnight at 4° C either alone or with graded dosages of the somatomedin standard or the plasmas. Membrane-bound somatomedin was determined by counting the centrifuged membrane pellet. The competitive effects of the three preparations were expressed as the percentage of controls containing only membrane and radioiodinated somatomedin. Each point is the mean ± SEM of data derived in seven separate assays. Reprinted with permission from Marshall, R. N., Underwood, L. E., Voina, S. J., Foushee, D. B., Van Wyk, J. J. 1974. *J. Clin. Endocrinol. Metab.* 39:283–92.

that the factors in serum of starved rats which inhibit sulfate incorporation in rat cartilage assays have no influence on the radioreceptor assay.

Yalow et al have developed a highly sensitive radioimmunoassay for somatomedin-B which can detect levels in plasma diluted 1:5000 or greater (30a). They have found that in some patients with acromegaly, definitive therapy leading to a decrease in growth hormone is accompanied by a fall in somatomedin-B. Preliminary findings suggest that the levels of somatomedin-B in plasma are in the microgram/milliliter range rather than in the nanogram range, which is common for most peptide hormones. Of even greater interest is that the species specificity of the assay is the same as that for growth hormone; i.e. the antibody cross-reacts with human and monkey plasma, but not with plasma from subprimate species.

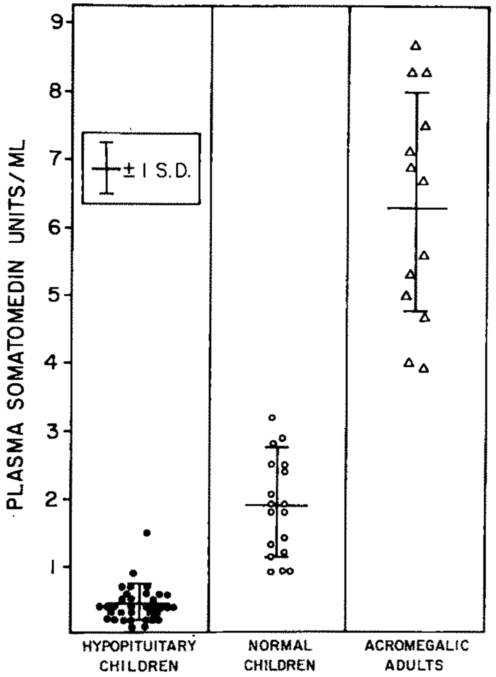


Figure 2 Plasma somatomedin levels determined by the competitive placental membrane binding assay in 35 hypopituitary children, 19 normal children, and 12 adults with acromegaly. Determinations were run in duplicate at plasma concentrations ranging between 1 μl/ml (in acromegalics) and 20 μl/ml (in hypopituitary children). The somatomedin activity in each sample was determined by comparison with the normal plasma pool standard. The mean somatomedin values ± 1 SD are indicated for each group.

INSULIN-LIKE EFFECTS OF SOMATOMEDIN

Biologic Effects

Somatomedin mimics the actions of insulin in all in vitro bioassay systems which measure insulin-like activity. In adipose tissue, it stimulates glucose oxidation and lipid synthesis and inhibits epinephrine-stimulated lipolysis (31–33). At subsaturating concentrations, the actions of somatomedin and insulin are additive, thus suggesting that the insulin-like effects of somatomedin and of insulin itself might operate through a common mechanism (32). In muscle, somatomedin stimulates membrane transport of sugar and amino acids and promotes the synthesis of protein and RNA (27, 34). In this system, the time course of the somatomedin effect is identical to that of insulin and different from that of growth hormone which is effective only after a period of preincubation (27, 35).

Insulin and somatomedin also affect adenylate cyclase in a similar manner. Under certain circumstances, insulin inhibits the epinephrine-induced rise of adenylate cyclase in adipose tissue and liver (36, 37). Tell et al found that somatomedin inhibited the rise in adenylate cyclase produced in lymphocytes and fat cells by epinephrine, in chondrocytes by parathyroid hormone, and in liver by prostaglandin (PGE_2) (38).

Interactions with Insulin Receptor

A rational explanation for these insulin-like effects of somatomedin was provided by the demonstration that in a wide variety of tissues, somatomedin is able to compete with ^{125}I-insulin for binding to the primary insulin receptor on the cell membrane (29, 32, 39). The potency ratio between somatomedin and insulin in insulin receptor assays is remarkably constant from tissue to tissue with one unit of somatomedin having the potency of approximately 100–350 μU of insulin. This potency ratio agrees very well with the biological potency ratios of the two hormones in the standard fat pad and rat diaphragm bioassays for insulin. Competitive binding assays for insulin-like activity are sensitive to immunoreactive insulin and other insulin-like substances as well as somatomedin. The insulin radioreceptor assay is therefore less specific than the radioreceptor assay for somatomedin-C (29). Using the radioreceptor assay for total insulin-like activity, we have been unable to detect differences in unextracted serum between hypopituitary, normal, and acromegalic patients.

Significance of Insulin-Like Effects of Somatomedin

It is now clear that the effects of somatomedin on carbohydrate and fat metabolism in assays for insulin-like activity are attainable only at relatively high somatomedin dosages. These effects probably have little physiologic significance in vivo. Conversely, the growth-stimulating effects of insulin, as measured in cartilage and tissue culture assays, are obtainable only with dosages of insulin far above the concentration of insulin in plasma.

The marked discrepancy between insulin: somatomedin potency ratios in growth-stimulating assays, as contrasted with assays based on insulin-like activity, may be explained by the existence of separate receptors for insulin and somatomedin on the same cell membranes. Membrane receptor assays utilizing ^{125}I-somatomedin-C are 100 times more sensitive to competition by unlabeled somatomedin than are similar assays using ^{125}I-insulin as the trace (29). Similarly, specific somatomedin receptor assays are markedly insensitive to insulin, since insulin concentrations several thousand fold higher than those found in plasma are required to observe a threshold effect (29). This insensitivity of the somatomedin receptor to insulin agrees remarkably well with the very large dosages of insulin required to mimic the effect of somatomedin in cell growth assays. The fact that somatomedin and insulin share overlapping biological activities and cross-react with their respective receptors, however, suggests that there may be some structural homology between the two hormones. It also suggests that the active sites on the molecule which bind to the insulin receptor probably differ from the active sites which interact with specific somatomedin receptor.

RELATIONSHIP BETWEEN SOMATOMEDIN AND NONSUPPRESSIBLE INSULIN-LIKE ACTIVITY

Less than 10% of the total plasma insulin-like activity, as measured in biological assays, can be neutralized by antibodies directed against pancreatic insulin (40). The remainder, which is immunologically distinct from insulin, is called nonsuppresible insulin-like activity (NSILA). Pancreatectomy does not diminish the content of NSILA, and NSILA levels are normal in diabetic patients (41, 42).

Froesch and his group in Zurich have pioneered in characterizing the chemical nature and biologic properties of NSILA. They have shown that the portion of NSILA which is soluble in cold acid ethanol is attributable to a peptide with a molecular weight of about 7500. The Zurich group has reported that this peptide, designated NSILA-S, is growth hormone dependent, that it stimulates sulfate incorporation in the rat cartilage assay, and that it competes for the insulin receptor in chick fibroblasts (43, 44). Furthermore, NSILA-S inhibits adenylate cyclase in fat cells (45) and is a mitogen for chick fibroblasts (46). Recently, Megyesi et al, using a preparation of NSILA-S furnished by Froesch, demonstrated that liver cell membranes possess a specific receptor for NSILA-S which is distinct from the insulin receptor. The potency ratios between insulin and NSILA-S in competing for the separate receptors are similar to the potency ratios between insulin and somatomedin-C in competing for the insulin and somatomedin-C receptors (47). Thus, evidence is rapidly accumulating that NSILA-S is either identical or at least only slightly different from somatomedin-C. It is unlikely that the issue of identity can be further resolved short of complete structural analysis of the two peptides.

RELATIONSHIP BETWEEN SOMATOMEDIN AND OTHER GROWTH FACTORS

A number of growth factors sharing some features in common with somatomedin have been described. Pierson & Temin have purified a small peptide from calf serum which stimulates the multiplication of fibroblasts in tissue culture (48). This peptide, which was given the designation "multiplication stimulating activity" (MSA), has activity in rat cartilage assays and in biological and competitive binding assays for insulin.

Dulak & Temin have shown that a specific clonal line of rat liver cells synthesizes sufficient quantities of MSA to permit exuberant proliferation in the absence of serum (49, 50). Several closely related peptides of approximately 7000 daltons have been isolated from the media in which these cells are grown. Samples of these peptides have proven highly active in our somatomedin-C receptor assay. Although growth hormone dependency has not been studied, MSA appears to be closely related to the somatomedins identified in human plasma.

Nerve growth factor (NGF), a peptide of about 12,000 daltons isolated from submandibular gland of male mice, triggers differentiative processes in the developing nervous system and maintains the mature nerve cell. NGF has a high degree of structural homology with proinsulin, and the two peptides have overlapping biological activities (51, 52). Whether NGF is influenced by growth hormone is not known.

Epidermal growth factor (EGF), a peptide of 6000 daltons, was also isolated from submandibular glands of male mice (53). This substance accelerates the eruption of the incisors and opening of the eyelids in young mice and stimulates epithelial cell proliferation in organ cultures. Furthermore, EGF can replace insulin in the growth of mammary gland explants (54) and stimulate DNA synthesis and cell replication in human fibroblasts (55). EGF is highly responsive to stimulation by androgenic hormones (56), but whether growth hormone plays any part in its regulation is not known.

Erythropoietin is a peptide growth factor specific for erythrocyte precursors and, like EGF and NGF, is responsive to androgens (57). It is also partially responsive to growth hormone since levels are high in acromegaly, low in hypopituitarism, and rise in response to growth hormone administration (58).

Gospodarowicz has recently isolated fibroblast growth factors from bovine brain and pituitary tissue (59). These substances, which are of about 13,000 daltons, appear to stimulate the same train of metabolic responses in fibroblast cultures as do somatomedins B and C, MSA, EGF, NSILA-S, and other less well characterized mitogenic peptides. The brain growth factor has the remarkable property of stimulating limb regeneration in the frog (59a). No information is yet available on whether this substance is subject to hormonal regulation.

The substances described above provide only a sampling of a lengthening list of mitogenic peptides which have in common the capacity to stimulate in their respective target tissues the set of biochemical reactions described by Hershko et al as a positive pleiotypic response (60). A rational scheme of classification and nomenclature will be possible only when the molecular structures of these substances are known and their hormonal control mechanisms delineated.

PURIFICATION AND CHEMICAL PROPERTIES OF THE SOMATOMEDINS

In contrast to the situation with other peptide hormones, no organ has been identified which contains a higher concentration of somatomedin than that present in native plasma. Progress in isolating and determining the chemical structure of these trace peptides has therefore required the fractionation of enormous quantities of outdated human plasma as starting material. Moreover, until very recently, the rate of progress has been limited by the slow and cumbersome bioassay methods needed to monitor purification.

In whole plasma, somatomedin activity is associated with proteins the size of albumin or larger (11, 61). After extraction of plasma with cold acid ethonal, 20–40% of the original biological activity is recovered in the soluble fraction and thus separated from the bulk of plasma proteins. Following gel chromatography of the acid ethanol extract in acid medium, somatomedin is recovered in fractions corresponding to peptides having a molecular weight between 5000 and 8000 (8, 11, 32, 62). Somatomedin is stable in acid solution and is not destroyed by boiling. It is inactivated by highly alkaline solutions and by treatment with β-mercaptoethanol. Thus, at least one disulfide linkage seems to be required for its biological activity.

Three separate somatomedins, provisionally designated by the suffixes A, B, and C, have thus far been isolated from acid ethanol extracts of Cohn fraction IV of outdated human plasma.[5] The Swedish group of Uthne, Hall, Sievertsson, and Fryklund have monitored their plasma fractionation by two entirely different bioassay systems, the embryonic chick cartilage sulfation factor assay (14) and a cell culture assay which depends on the stimulation of thymidine uptake by human glial-like cells (27). Somatomedin-A is their designation for a neutral peptide of about 7000 daltons which is most active in chick cartilage. Somatomedin-B is an acidic peptide of about 4500 daltons which is most active in glial cells.

Using somewhat different methods, we have identified two peptide fractions with activity in the rat cartilage assay. After isoelectric focusing in 6-M urea, a majority of the biological activity focused in the basic region (pH 8.4–9.4), whereas less than one third of this activity could be accounted for in the neutral region. Conversely, the neutral fractions were more active than the basic fractions in competitive binding assays for insulin-like activity. The basic peptide, designated somatomedin-C, was subsequently isolated in essentially pure form by preparative acrylamide gel electrophoresis and a series of additional chromatographic steps (4). In addition to its activity in skeletal tissues, somatomedin-C is highly active in stimulating DNA synthesis and the rate of mitosis in chick fibroblast cultures.

The insulin-like activity of our basic fraction and somatomedin activity of our neutral fraction proved to be intrinsic properties of discrete peptides rather than contaminants, since the respective somatomedin:insulin ratios remained constant through subsequent purification procedures. Since, as discussed in previous sections, it seems likely that different portions of the molecule are responsible for these activities, it is possible that the neutral and basic peptides represent chemical variants of a single naturally occurring peptide. Since these forms are similar in size ∼ 7000 daltons), it is conceivable that alteration of charge properties during the course of purification might selectively alter that portion of the molecule which interacts with the insulin or somatomedin receptors. Until complete structural comparisons are available between the several peptides isolated in different laboratories, it is our tentative conclusion that somatomedins A and C and NSILA-S are derived from a single naturally occurring peptide, and that this peptide accounts for most of the "sulfation factor" activity of native plasma. Somatomedin-B, however, appears to represent a quite different substance.

PRODUCTION AND DISPOSAL OF SOMATOMEDIN

Since somatomedin does not appear to be stored in any organ, it may be inferred that this substance is not manufactured in a classic endocrine tissue possessing secretory granules. There is now growing evidence that the liver is a major site of somatomedin generation. Following the intravenous administration of ^{125}I-growth hormone, a high percentage of the radioactive label is taken up by hepatic cells (63). Somatomedin activity in serum declines with partial hepatectomy and rises again as the liver regenerates (64). Lastly, perfusion of isolated rat liver with growth hormone results in a prompt increase of somatomedin activity in the effluent (65). Fragmentary evidence also suggests that muscle and kidney may also be sites of somatomedin generation (66, 67). Other sites of somatomedin generation have not been excluded, and studies so far have left unsettled the question of whether somatomedin is synthesized de novo or is a cleavage product of growth hormone itself.

---

[5] We are indebted to Professor Bertil Åberg, M.D., Vice-President and Director of Research, AB Kabi Laboratories, Stockholm, and Knut Uthne, Ph.D., of the Kabi Laboratories for supplying us with the acid ethanol extracts from which purification of somatomedin was accomplished.

In comparison to most peptide hormones, production and disposal of somatomedin appears to be relatively slow. Rises in plasma somatomedin levels are not observed until 3 hr after the intravenous injection of human growth hormone (hGH) into hypopituitary children (68). By the i.m. route, growth hormone administration for several days is often required to detect substantial increases (69, 70). Following hypophysectomy in the rat, somatomedin disappears from plasma with a half time of 3–4 hr (71).

The long half-life of somatomedin may be due to its association in native plasma with a macromolecular binding protein. Hintz et al have shown that at pH 8.1 the plasma somatomedin activity is associated with plasma protein fractions $\geq$50,000 daltons, but under acid conditions this activity is readily dissociated and recovered after gel chromatography with peptides in the molecular weight range of 5000–9000. At neutral pH, "small" somatomedin readily recombines with the larger carrier fractions (72). Using our preparation of ^{125}I-somatomedin-C, Marshall has obtained electrophoretic evidence that the major somatomedin carrier protein is albumin (unpublished observations).

SOMATOMEDIN LEVELS IN PLASMA

Somatomedin Levels in Normal, Hypopituitary, and Acromegalic Subjects

Accurate measurement of the somatomedin content of plasma has been difficult due to the inherent imprecision of the cumbersome bioassay techniques, the presence of variable quantities of inhibitory substances, and the absence of a uniform international standard. The somatomedin unit is defined as the biological activity in 1 ml pooled serum from normal adult males. The potency of such standards varies from laboratory to laboratory and even from time to time within the same laboratory. The recently developed competitive membrane binding assay for somatomedin shows promise of eliminating many of the difficulties encountered with bioassays, and the advent of radioimmunoassays for somatomedin should further resolve these problems.

A modest rise in plasma somatomedin levels has been reported with increasing age from birth to adulthood (19, 73). While there are no known sex differences in somatomedin levels and there appear to be no circadian rhythms (19, 74), final judgment on such phenomena must await more precise methods of measurement. Bioassayable levels in hypopituitary animals and humans generally average about one third to one fourth of those of normal subjects. Partial or even complete restoration of somatomedin levels to normal is achieved by growth hormone treatment (19, 69). When measured by bioassay, levels in acromegalic patients are somewhat variable (62, 73, 74). By using the somatomedin-C competitive membrane binding assay, we have found high plasma levels in 12 successive patients with active acromegaly (Figure 2).

Possibly because somatomedin levels are more stable they have correlated better with growth rate than have plasma growth hormone levels. Hall & Olin have shown that the initial, rapid phase of growth which occurred with the institution of hGH therapy in the hypopituitary child was paralleled by increases in plasma somatomedin levels (75). With prolonged treatment, however, the somatomedin levels paralleled the decline in growth rate.

Conditions Characterized by a Paradoxical Relationship between Growth Hormone and Somatomedin Levels

A group of conditions have been described in which somatomedin levels are inversely related to growth hormone levels. These conditions are potentially important because of the light they may shed on the relationship between growth hormone and somatomedin.

Laron et al have described a unique form of familial dwarfism occurring in Asiatic Jews. Affected patients bear a striking clinical resemblance to children with severe, idiopathic hypopituitary dwarfism (76, 77). These children, unlike hypopituitary children, exhibit striking elevations in plasma growth hormone levels and exaggerated responses to the usual stimuli for growth hormone release. Their somatomedin levels are low and do not rise when growth hormone is administered parenterally (69, 78, 79). These children also fail to respond metabolically or to grow more rapidly when given exogenous growth hormone (78, 80). It is now generally accepted that their basic defect is resistance to growth hormone in peripheral tissues.

In kwashiorkor, several investigators have reported that growth hormone levels are strikingly elevated (81, 82), and somatomedin levels are depressed (19, 83). In these patients, there is apparently an acquired reversible defect in somatomedin generation. With refeeding, growth hormone levels fall while somatomedin levels rise (73, 84). Van den Brande has reported that the levels of somatomedin revealed by bioassay in such patients may be spuriously low due to the presence of inhibitory substances (19). It is of interest to speculate on the survival value of a mechanism which might increase the lipolytic, protein-sparing actions of growth hormone while at the same time preventing the squandering of protein in growth.

Another apparent paradox between growth hormone and somatomedin is seen in some patients who exhibit normal or accelerated growth after surgical removal of a craniopharyngioma (85, 86). These children usually have undetectable levels of growth hormone but normal somatomedin levels (87, 88). Patients with this syndrome typically have severe hyperphagia and massive obesity. The stimulus for somatomedin generation in these patients is unknown.

Somatomedin in Chronic Renal Insufficiency

Growth failure is common in patients with chronic renal insufficiency and may be due, in part, to somatomedin deficiency. Saenger et al (89) have reported low somatomedin levels by bioassay in uremic patients. These levels rose to normal in those patients who had substantial improvement in creatinine clearance and growth rate following transplantation. Measurements of somatomedin in uremic patients by the sulfation factor bioassay may be spuriously low unless corrections are made for the high serum sulfate levels which may be present.

Interaction of Somatomedin with Other Hormones

Sustained high plasma glucocorticoid levels, of either endogenous or exogenous origin, may produce growth arrest by several mechanisms. Diminished secretion of growth hormone plays a role in only a few of such patients (90, 91), whereas somatomedin levels are regularly depressed (21). Glucocorticoids also attenuate the growth hormone induced rise of somatomedin in hypophysectomized animals (21). While the primary effect of glucocorticoids seems to be the inhibition of somatomedin production, there is also evidence that at higher dosages, glucocorticoid acts peripherally to inhibit the action of somatomedin on cartilage (6, 22, 92).

Large doses of estrogens inhibit growth in young animals (93, 94), counteract many of the metabolic effects of growth hormone in acromegaly (95), and depress somatomedin levels in plasma of humans (96). Since in vitro addition of estrogen to somatomedin assays does not inhibit the effects of somatomedin on cartilage (29, 96), it is probable that the effects of estrogen on growth are mediated by virtue of its depressive effect on somatomedin generation.

CONCLUSION

Although the existence of tissue-specific growth factors now seems well established, this field of research is yet in its infancy and caution must be observed in drawing conclusions concerning the physiologic significance of these substances in living organisms. Virtually nothing is known of the mechanisms involved in their formation nor of the factors governing their homeostatic regulation. Sufficient quantities of the purified peptides are not yet available to study the in vivo effects of these substances following acute and chronic administration. It is apparent, however, that further knowledge about a group of peptides which stimulate the growth of specific tissues or groups of tissues potentially could fill large gaps in our present understanding of growth and differentiation and conceivably could have therapeutic applications of inestimable value.

Literature Cited

1. Salmon, W. D. Jr., Daughaday, W. H. 1957. *J. Lab. Clin. Med.* 49:825–36
2. Daughaday, W. H. et al 1972. *Nature* 235:107
3. Hall, K., Van Wyk, J. J. 1974. *Current Topics in Experimental Endocrinology*, ed. V. H. T. James, L. Martini, 155–78. New York: Academic
4. Van Wyk, J. J. et al 1974. *Recent Prog. Horm. Res.* 30:259–318
5. Daughaday, W. H., Reeder, C. 1966. *J. Lab. Clin. Med.* 68:357–68
6. Daughaday, W. H., Mariz, I. K. 1962. *J. Lab. Clin. Med.* 59:741–52
7. Salmon, W. D. Jr., DuVall, M. R. 1970. *Endocrinology* 86:721–27
8. Van Wyk, J. J., Hall, K., Weaver, R. P. 1969. *Biochim. Biophys. Acta* 192: 560–62
9. Raben, M. S., Murakawa, S., Matute, M. 1972. *Growth and Growth Hormone*, ed. A. Pecile, E. E. Muller, 124–31. Amsterdam: Excerpta Medica
10. Van den Brande, J. L., Van Wyk, J. J.,

Weaver, R. P., Mayberry, H. E. 1971. *Acta Endocrinol.* 66:65–81

11. Van Wyk, J. J., Hall, K., Van den Brande, J. L., Weaver, R. P. 1971. *J. Clin. Endocrinol. Metab.* 32:389–403
12. Van Wyk, J. J. et al 1972. See Ref. 9, pp. 155–67
13. Yde, H. 1968. *Acta Endocrinol.* 57: 557–64
14. Hall, K. 1970. *Acta Endocrinol.* 63: 338–50
15. Van den Brande, J. L., DuCaju, M.V.L. 1974. *Acta Endocrinol.* 75:233–42
16. Phillips, L. S., Herington, A. C., Daughaday, W. H. 1974. *Endocrinology* 94:856–63
17. Salmon, W. D. Jr. 1972. See Ref. 9, p. 180
18. Salmon, W. D. Jr. 1974. *Advances in Human Growth Hormone Research,* ed. S. Raiti, 76–94. Washington, DC: DHEW Publ. No. NIH–74–612
19. Van den Brande, J. L., DuCaju, M. V. L. 1974. See Ref. 18, pp. 98–115
20. Delcher, H. K., Eisenbarth, G. S., Lebovitz, H. E. 1973. *J. Biol. Chem.* 248:1901–5
21. Phillips, L. S., Herington, A. C., Daughaday, W. H. 1974. See Ref. 18, pp. 50–67
22. Elders, M. J., Wingfield, B. S., McNatt, L. M., Hughes, E. R. 1973. *Pediat. Res.* 7:324 (Abstr.)
23. Holley, R. W., Kiernan, J. A. 1971. *Growth Control in Cell Culture,* ed. G. E. W. Wolstenholme, J. Knight, 3–10. London: Churchill—Livingstone
24. Temin, H. M., Pierson, R. W. Jr., Dulak, N. C. 1972. *Growth, Nutrition and Metabolism in Cells in Culture,* 50–81. New York: Academic
25. Gey, G. O., Thalhimer, W. 1924. *J. Am. Med. Assoc.* 82:1609
26. Salmon, W. D. Jr., Hosse, B. R. 1971. *Proc. Soc. Exp. Biol. Med.* 136:805–8
27. Uthne, K. 1973. *Acta Endocrinol. Suppl.* 175:1–35
28. Wasteson, A., Uthne, K., Westermark, B. 1973. *Biochem. J.* 136:1069–74
29. Marshall, R. N., Underwood, L. E., Voina, S. J., Foushee, D. B., Van Wyk, J. J. 1974. *J. Clin. Endocrinol. Metab.* 39:283–92
30. Van Wyk, J. J., Underwood, L. E., Hintz, R. L., Voina, S. J., Weaver, R. P. 1974. See Ref. 18, p. 25

30a. Yalow, R., Hall, K., Luft, R. 1975. *J. Clin. Invest.* 55:127–37

31. Hall, K., Uthne, K. 1971. *Acta Med. Scand.* 190:137–43
32. Clemmons, D. R., Hintz, R. L., Underwood, L. E., Van Wyk, J. J. 1974. *Isr. J. Med. Sci.* 10:1254–62
33. Underwood, L. E., Hintz, R. L., Voina, S. J., Van Wyk, J. J. 1972. *J. Clin. Endocrinol. Metab.* 35:194–98
34. Salmon, W. D. Jr., DuVall, M. R. 1970. *Endocrinology* 87:1168–80
35. Kostyo, J. L., Nutting, D. F. 1973. *Horm. Metab. Res.* 5:167–72
36. Illiano, G., Cuatrecasas, P. 1972. *Science* 175:906–8
37. Hepp, K. D., Renner, R. 1973. *Fed. Eur. Biochem. Soc. Lett.* 20:191–94
38. Tell, G. P. E., Cautrecasas, P., Van Wyk, J. J., Hintz, R. L. 1973. *Science* 180:312–15
39. Hintz, R. L., Clemmons, D. R. Underwood, L. E., Van Wyk, J. J. 1972. *Proc. Nat. Acad. Sci. USA* 69:2351–53
40. Oelz, O., Froesch, E. R., Bunzli, H. F., Humbel, R. E., Ritschard, W. J. 1972. *Handbook of Physiology—Endocrinology 1,* 685–702. Bethesda: Am. Physiol. Soc.
41. Steinke, J., Sirek, A., Lauris, V., Lukens, F. D. W., Renold, A. E. 1962. *J. Clin. Invest.* 41:1699–1707
42. Froesch, E. R. et al 1967. *Recent Progr. Horm. Res.* 23:565–616
43. Zingg, A. E., Froesch, E. R. 1973. *Diabetologia* 9:472–76
44. Froesch, E. R. et al 1974. *Acta Endocrinol. Suppl.* 184:183–84 (Abstr.)
45. Renner, R., Hepp, K. D., Humbel, R. E., Froesch, E. R. 1973. *Horm. Metab. Res.* 5:56–57
46. Morell, B., Froesch, E. R. 1973. *Eur. J. Clin. Invest.* 3:119–23
47. Megyesi, K. et al 1974. *Biochem. Biophys. Res. Comm.* 57:307–15
48. Pierson, R. W. Jr., Temin, H. M. 1972. *J. Cell. Physiol.* 79:319–29
49. Dulak, N. C., Temin, H. M. 1973. *J. Cell. Physiol.* 81:153–60
50. Dulak, N. C., Temin, H. M. 1973. *J. Cell Physiol.* 81:161–70
51. Angeletti, R. H., Bradshaw, R. A. 1971. *Proc. Nat. Acad. Sci. USA* 68:2417–20
52. Frazier, W. A., Angeletti, R. H., Bradshaw, R. A. 1972. *Science* 176:482–88
53. Cohen, S. 1962. *J. Biol. Chem.* 237:1555–62
54. Turkington, R. W. 1969. *Exp. Cell Res.* 57:79–85
55. Hollenberg, M. D., Cautrecasas, P. 1973. *Proc. Nat. Acad. Sci.* 70:2964–68
56. Barthe, P. L., Bullock, L. P., Mowszowicz, I., Bardin, C. W., Orth, D. N. 1974. *Endocrinology* 95:1019–25
57. Jepson, J. H. McGarry, E. E., Lowenstein, L. 1968. *Arch. Intern. Med.* 122:265–70
58. Jepson, J. H. McGarry, E. E., 1972. *Blood* 39:238–48

59. Gospodarowicz, D. et al 1974. *Nature* 249:123–27
59a. Gospodarowicz, D. et al 1974. *Nature*. In press
60. Hershko, A., Mamont, P., Shields, R., Tomkins, G. M. 1971. *Nature New Biol.* 232:206–11
61. Bala, R. M., Ferguson, K. A., Beck, J. C. 1970. *Endocrinology* 87:506–16
62. Hall, K. 1972. *Acta Endocrinol. Suppl.* 163:1–52
63. Mayberry, H. E., Van den Brande, J. L., Van Wyk, J. J., Waddell, W. J. 1971. *Endocrinology* 88:1309–17
64. Uthne, K., Uthne, T. 1972. *Acta Endocrinol.* 71:255–64
65. McConaghey, P., Sledge, C. B. 1970. *Nature* 225:1249–50
66. Hall, K., Bozovic, M. 1969. *Horm. Metab. Res.* 1:235–40
67. McConaghey, P., Dehnel, J. 1972. *J. Endocrinol.* 52:587–88
68. Hall, K. 1971. *Acta Endocrinol.* 66:491–97
69. Daughaday, W. H., Laron, Z., Pertzelan, A., Heins, J. N. 1969. *Trans. Assoc. Am. Physicians* 82:129–38
70. Van den Brande, J. L. 1973. *Plasma somatomedin: Studies on some of its characteristics and on its relationship with growth hormone.* PhD thesis. Erasmus Universiteit te Rotterdam, The Netherlands. 117 pp.
71. Daughaday, W. H. 1971. *Advan. Intern. Med.* 17:237–63
72. Hintz, R. L., Orsini, E. M., Van Camp, M. G. 1974. *56th Meet. Endocrine Soc., Atlanta, Ga.*, Abstr. 31
73. Almqvist, S., Rune, I. 1961. *Acta Endocrinol.* 36:566–76
74. Daughaday, W. H., Salmon, W. D. Jr., Alexander, F. 1959. *J. Clin. Endocrinol. Metab.* 19:743–58
75. Hall, K., Olin, P. 1972. *Acta Endocrinol.* 69:417–33
76. Laron, Z., Pertzelan, A., Mannheimer, S. 1966. *Isr. J. Med. Sci.* 2:152–55
77. Laron, Z. et al 1972. See Ref. 9, pp. 458–82
78. New, M. I., Schwartz, E., Parks, G. A., Landey, S., Wiedemann, E. 1972. *J. Pediat.* 80:620–26
79. Van den Brande, J. L. et al 1974. *Arch. Dis. Childhood* 49:297–304
80. Laron, Z., Pertzelan, A., Karp, M., Kowadlo-Silbergeld, A., Daughaday, W. H. 1971. *J. Clin. Endocrinol. Metab.* 33:332–42
81. Pimstone, B. L., Wittmann, W., Hansen, J. D. L., Murray, P. 1966. *Lancet* 2:779–80
82. Beas, F., Muzzo, S. 1973. *Endocrine Aspects of Malnutrition*, ed. L. I. Gardner, P. Amacher, 1–18. Santa Ynez, California: The Kroc Foundation
83. Grant, D. B., Hambley, J., Becker, D., Pimstone, B. L. 1973. *Arch. Dis. Childhood* 48:596–600
84. Pimstone, B. L., Barbezat, G., Hansen, J. D. L., Murray, P. 1968. *Am. J. Clin. Nutr.* 21:482–87
85. Frasier, S. D., Smith, F. G. 1968. *Am. J. Dis. Child.* 116:311–14
86. Kenny, F. M. et al 1968. *J. Pediat.* 72:766–75
87. Finkelstein, J. W., Kream, J., Ludan, A., Hellman, L. 1972. *J. Clin. Endocrinol. Metab.* 35:13–17
88. Kenny, F. M., Guyda, H. J., Wright, J. C., Friesen, H. G. 1973. *J. Clin. Endocrinol. Metab.* 36:378–80
89. Saenger, P. et al 1974. *Pediat. Res.* 8:163–69
90. Strickland, A. L., Underwood, L. E., Voina, S. J., French, F. S., Van Wyk, J. J. 1972. *Am. J. Dis. Child.* 123:207–13
91. Morris, H. G., Jorgensen, J. R., Jenkins, S. A. 1968. *J. Clin. Invest.* 47:427–35
92. Peck, W. A., Brandt, J., Miller, I. 1967. *Proc. Nat. Acad. Sci. USA* 57:1599–1606
93. Zondek, E. 1936. *Lancet* 2:842–47
95. Schwartz, E., Echemendia, E., Schiffer, M., Panariello, V. A. 1969. *J. Clin. Invest.* 48:260–70
96. Weidemann, E., Schwarz, E. 1972. *J. Clin. Endocrinol. Metab.* 34:51–58

Copyright 1975. All rights reserved.

ACTIVE FORM OF THE THYROID HORMONE

❖7164

Sidney H. Ingbar, M.D.
Thorndike Laboratory of the Harvard Medical School at the Beth Israel Hospital, and the Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115

Lewis E. Braverman, M.D.
St. Elizabeth's Hospital, and the Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts 02111

INTRODUCTION

It is increasingly evident that the chemical form in which many hormones are active at the tissue level differs from that in which the hormone is either synthesized in or secreted from its gland of origin. Many peptide hormones, for example, are synthesized as large-molecular "prohormones" that, either prior to or following their secretion into the blood, are modified to yield smaller moieties of greater biological potency. Although not previously viewed in the context of such relationships, an analogous situation pertains in the case of the thyroid hormones. Secretion of the hormonally active iodoamino acids requires proteolytic cleavage of the prohormone precursor, thyroglobulin. The further question which now arises is whether the major secretory product of the thyroid gland, thyroxine (T_4), has intrinsic biological activity or is itself merely a prohormone from which an ultimately active form is peripherally derived. This brief review will describe both the recent findings which have raised this question and some of the physiological implications that flow therefrom. The clinical correlates of these findings have been thoroughly reviewed by Utiger in the preceding issue of this volume (1).

T_4 was originally thought to be both the sole hormone secreted by the thyroid gland and the direct initiator of thyroid hormone action at the cellular level. In the early 1950s, however, 3,5,3'-triiodothyronine (T_3) was identified as a constituent of thyroid tissue, a secretory product of the gland, and a compound with several times the overall metabolic potency of T_4 in both animals and man (2–4). These findings raised questions as to whether and to what extent T_4 is converted to T_3 in peripheral tissues and whether such conversion could be a prerequisite to the metabolic action of T_4. Although several studies thereafter revealed the conversion of T_4 to T_3 in intact animals or tissues in vitro (5–9), early efforts to demonstrate this process in man were inconclusive (10–11). In 1970, however, substantial concentrations of both stable and radioiodine-labeled T_3 were found in the blood of athyreotic patients receiving either oral or parenteral maintenance therapy with mixtures of highly purified stable and ^{125}I-labeled T_4 (12). This clear demonstration of the operation of pathways of T_4 to T_3 conversion in man, since repeatedly confirmed (13, 14), provided the main stimulus to current interest in the nature of the active form of the thyroid hormone and related questions. Figure 1 illustrates the structural formulae of T_4 and those of its peripheral metabolites herein discussed.

QUANTITATION OF T_4 TO T_3 CONVERSION

The question which next arises concerns the extent to which the process of peripheral T_4 to T_3 conversion operates in normal and abnormal states. Two general methods for estimating this function are available. The first is the so-called replacement method, which can be employed only in patients whose thyroid function is negligible and who are receiving physiological replacement doses of levothyroxine

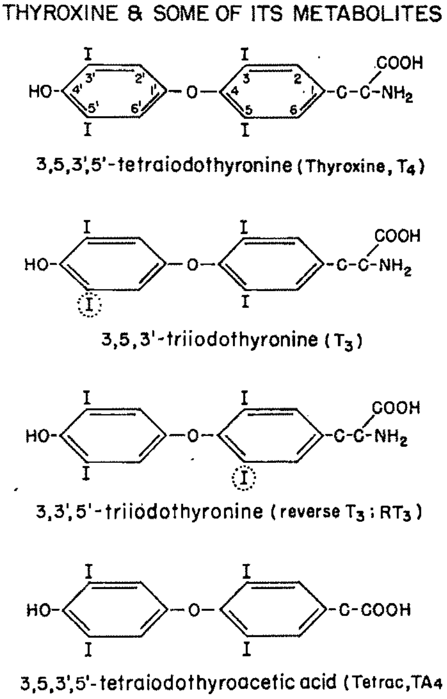

THYROXINE & SOME OF ITS METABOLITES
I
I
HO
3'
2'
4'
1'
5'
6'
O
3
2
4
1
5
6
C
COOH
C
NH₂
I
I
3,5,3',5'-tetraiodothyronine (Thyroxine, T₄)
3,5,3'-triiodothyronine (T₃)
3,3',5'-triiodothyronine (reverse T₃ ; RT₃)
C-COOH
3,5,3',5'-tetraiodothyroacetic acid (Tetrac, TA4)

Figure 1 Structural formulae of thyroxine (T_4) and several of its metabolites.

(14, 15). Hence, it is useful only in patients who are either hypothyroid or euthyroid with suppressible thyroid function. Differentially labeled T_4 and T_3 are used to measure the metabolic clearance rate (MCR) of each hormone. The quantity of each hormone metabolized daily is then calculated as the product of its MCR and plasma concentration. If a steady state exists, daily T_4 turnover reflects the quantity of T_4 provided by the exogenous medication, while the quantity of T_3 degraded closely approximates the rate at which T_3 is produced by the monodeiodination of T_4. The latter is a "blood production rate," however, since any T_3 generated from T_4 that undergoes further degradation at its site of formation without first entering the blood will not be measured. The magnitude of this error, which is also operative in the alternate method described below, is uncertain.

The second method also utilizes differentially labeled T_4 and T_3, together with measurements of stable serum concentrations, to define MCRs and total rates of disposal for the two hormones. In addition, the proportion of T_4 metabolized via processes that generate T_3 can be calculated from the MCRs of the two hormones and the rate of appearance in blood of labeled T_3 derived from the administered labeled T_4 (13). This is determined by chromatographic separation and analysis of sequential serum samples. Although time consuming, technically difficult, and subject to multiple sources of experimental error, this "isotopic ratio" method has several advantages over the replacement method. First, it can be employed in the presence of a functioning thyroid, at whatever level of function pertains. Second, since estimates of both total T_3 disposal and the quantity of T_3 generated peripherally from T_4 are obtained, the contribution of thyroid secretion to total T_3 production can be calculated as the difference between the two.

Rather surprisingly, in view of the potential errors in the isotopic ratio method, studies by the two techniques in separate groups of euthyroid individuals or patients with treated hypothyroidism have agreed rather well and have indicated that in such patients approximately 35% of the daily turnover of T_4 occurs via pathways that lead to the generation of T_3 (13–15).

SECRETION OF T_3 BY THE NORMAL THYROID

A secondary question that arises is the extent to which peripheral conversion of T_4 and direct thyroid secretion each contribute to the normal daily production of T_3 (16). Since approximately 80 μg of T_4 are degraded daily, a 35% rate of conversion to T_3 would suggest that approximately 24 μg of T_3 are derived daily from T_4, account being taken of the lesser molecular weight of T_3. Since the total production rate of T_3 is estimated to be about 28 μg per day (17), the fraction of T_3 production provided by direct secretion would be about 15%. It should be noted, however, that these values are derived from multiple measurements, all of which are subject to experimental error. Particularly this is true in the case of analyses of stable serum T_3, which vary widely among different laboratories. Hence, it is possible either that this fraction is underestimated, or that peripheral conversion of T_4 accounts for essentially all of T_3 production. Evidence in favor of the latter view is provided by the high T_4/T_3 ratio (approximately 20/1) in the thyroglobulin of the normal human thyroid (18), and by studies of serum T_4 and T_3 concentrations in hypothyroid patients receiving levothyroxine therapy. When doses are adjusted so that serum T_4 concentration is equal to that in a matched group of normals, serum T_3 concentrations in the two groups are also virtually the same (19). Since the thyroid secretory rate of the patients with treated hypothyroidism should be virtually nil, the findings imply a comparatively small secretory rate for T_3 in the normal individual. The doses of levothyroxine required to establish a normal serum T_4 concentration (0.15–0.20 mg daily) are substantially lower than those formerly recommended, and since such doses are also sufficient to normalize the serum T_3, it is apparent that neither the higher doses formerly recommended nor combinations of T_4 and T_3 are required for treatment.

T_4 AS A PROHORMONE

Since T_4 is converted to T_3 and T_3 is more active metabolically than T_4, it would be expected that any factor that reduces conversion of T_4 to T_3 would decrease the metabolic effectiveness of T_4. This postulate is supported by strong experimental evidence. Propylthiouracil (PTU) is known to decrease the overall rate of T_4 deiodination in animals and man (20–24). In experiments derived from this observation it has been shown that in hypothyroid rats and humans maintained with replacement doses of T_4, the serum concentration of T_3 declines when PTU is administered, indicating that peripheral generation of T_3 has been inhibited. Associated with this inhibition are evidences of decreased hormonal activity, including increases in both basal serum thyroid-stimulating hormone (TSH) and responsiveness to thyrotropin-releasing hormone (TRH), and, in animals, decreases in both growth rate and the activity of certain thyroid hormone-induced enzymes (25–28).

Such experiments clearly demonstrate that conversion of T_4 to T_3 results in amplification of the metabolic activity of the original molecule. In this sense, T_4 can be considered a prohormone, without, however, implications concerning the presence or absence of intrinsic activity within the T_4 molecule itself.

INTRINSIC BIOLOGICAL ACTIVITY OF T_4

The fact that T_4 is a prohormone for T_3 leads inevitably to the question of whether T_4 itself has intrinsic biologic potency. Some have suggested that since approximately one third of T_4 is converted to T_3, and T_3 has approximately three times the metabolic potency of T_4, all of the metabolic activity of T_4 can be accounted for by the T_3 generated from it and none need be ascribed to T_4, per se (25, 29). This postulate is currently under active investigation.

Although the effects of T_4 and T_3 on a great many in vitro biological systems have been studied, the data provide little direct evidence concerning the intrinsic biological potency of T_4. For such observations to contribute meaningfully to this question, several requirements must be met. First, relative potencies of T_4 and T_3 in the system under study must be determined. Next, in the light of this information, careful examinations must be made to determine whether any apparent activity of T_4 could be due to T_3, either present as an initial contaminant or arising from the deiodination of T_4 during the course of the reaction being observed. Finally, there must be some assurance that the in vitro effect induced bears close relationship to an action of thyroid hormone in vivo, i.e. that the response observed truly has physiological relevance. We have been unable to find reports of studies that conform to the foregoing criteria.

Observations in a variety of clinical states, however, strongly indicate that T_4 does have intrinsic biological activity. These are mainly states in which the clinical or physiological condition conforms more closely to the serum concentration of T_4, or to the balance between the concentrations of T_4 and T_3, than to the concentration of T_3 itself. First, in a moderate proportion of patients with untreated hypothyroidism, values for the serum total and free T_3 concentrations are within the normal range, while serum TSH remains elevated and serum T_4 concentration is decreased (30). Second, there are several situations in which serum T_3 concentration is clearly *above* normal, while serum TSH is increased. This is seen in some clinically euthyroid patients with mild thyroid failure in whom the serum T_4 concentration is within the normal range (31), as well as in some cases of severe endemic iodine deficiency in whom the serum T_4 concentration is low (32, 33). Moreover, sequential measurements of serum T_3 concentration in hypothyroid patients rendered euthyroid by single daily doses of T_3 reveal that the serum T_3 concentration integrated over any 24-hr period is substantially in excess of normal (16, 30).

There are, in addition, several situations in which subnormal serum concentrations of both total and free T_3, in association with normal concentrations of T_4, are accompanied by either no evidence or very slight evidence that peripheral supplies of hormone are inadequate. In patients subjected to total caloric restriction for several weeks, serum total and free T_3 concentrations usually decline into the hypothyroid range, while serum T_4 concentration is essentially unchanged (34). Here, both basal serum TSH and the TSH response to TRH are normal. Similarly, patients with severe systemic illness or hepatic cirrhosis may display greatly depressed values for serum total and free T_3, normal serum T_4 concentrations, and either normal or only slightly elevated values for serum TSH concentration (35–38).

If, as seems likely, most of the circulating T_3 is derived peripherally from the deiodination of T_4, then the decrease in serum T_3 that occurs in the foregoing conditions most likely reflects a decreased activity of this process. For the present, it is unclear why a decrease in T_3 generation induced by PTU is associated with pituitary activation, while the decrease that accompanies starvation or illness is not. Be that as it may, all of the foregoing observations indicate that neither the level of TSH secretion, which is considered the most sensitive reflection of alterations in peripheral hormone supply, nor the clinical state of the patient, correlate perfectly with the serum T_3 concentration, as would be expected if T_3 were the only active hormone.

OTHER METABOLITES OF T_4

In normal humans, approximately 20% of the daily disposal of T_4 takes place through fecal excretion, mainly of glucuronyl and sulfonconjugates. Possibly 5–10%

is metabolized via deamination and oxidative decarboxylation of the alanine side chain to yield tetraiodothyroacetic acid, a compound of low biologic potency (12).

The remaining approximately 70% of T_4 is degraded by deiodinative processes, and of this roughly half seems to undergo initial monodeiodination in the outer or phenolic ring to yield T_3. It is tempting to raise the possibility that the remaining half may initially undergo monodeiodination of the inner ring to yield 3,3',5'-T_3, or "reverse" T_3 (RT_3). The latter pathway has not been conclusively demonstrated to operate in man, but RT_3 has been found in normal human serum by radioimmunoassay (39). The possibility that the initial deiodination of T_4 may give rise to equal quantities of T_3 and RT_3 has led to the postulate that the enzymes which deiodinate T_4 do so randomly, initially removing one iodine atom from either the inner or outer ring nonpreferentially (14, 25). A corollary of this theory of random deiodination is that the quantity of T_3 formed from T_4 would always be half of the quantity of T_4 undergoing deiodination, and that no physiological, pathophysiological, or pharmacological factor could alter this proportion.

A priori, one would be surprised if such an important process were not subject to physiological regulation, and indeed substantial evidence suggests that it is. Studies in the dog have revealed that labeled T_3 is generated from labeled T_4 when the animal is intact, but that labeled RT_3 becomes a prominent metabolite of T_4 when the animal is hepatectomized (5). This suggests that monodeiodination of the outer ring occurs mainly in the liver and monodeiodination of the inner ring mainly in the extrahepatic tissues. In addition, the studies in human starvation alluded to above revealed that although generation of T_3 from T_4 was decreased, overall clearance and deiodination of T_4 was unaltered. Clearly, the proportion of T_4 deiodinated that gave rise to T_3 must have been decreased. Finally, although serum T_4 concentration in human umbilical cord blood is slightly elevated, owing to the increase in thyroxine-binding globulin, values for the serum T_3 concentration are well below those found in normal adults (40–42), while concentrations of RT_3 appear unusually high (39). This suggests a diversion of deiodinative pathways from those that generate T_3 to those that generate RT_3.

When administered systemically, RT_3 has little or no metabolic potency, but it is uncertain whether this would be true of any RT_3 generated within the cell. If, however, RT_3 is truly inactive, or is in fact an inhibitor of T_3 as has been suggested (43), then modulation of the alternate pathways of T_4 metabolism could provide an effective means for modifying the balance between activation and inactivation of T_4.

ACKNOWLEDGMENT

The studies described herein that were carried out in the authors' laboratories were supported in part by Grant No. AM-16497 and Grant No. AM-07917 from the National Institute of Arthritis, Metabolism and Digestive Diseases, National Institutes of Health, Bethesda, Maryland.

Literature Cited

1. Utiger, R. D. 1974. *Ann. Rev. Med.* 25:41–53
2. Gross, J., Pitt-Rivers, R. 1952. *Lancet* 1:439–41
3. Gross, J., Pitt-Rivers, R., Trotter, W. R. 1952. *Lancet* 1:1044–45
4. Lerman, J. 1952. *J. Clin. Endocrinol. Metab.* 13:1341–46
5. Flock, E. V., Bollman, J. L., Grindlay, J. H., Stobie, G. H. 1961. *Endocrinology* 69:626–37
6. Albright, E. C., Larson, F. C. 1959. *J. Clin. Invest.* 38:1899–1903
7. Gross, J., Leblond, C. P. 1951. *Proc. Soc. Exp. Biol. Med.* 76:686–89
8. Albright, E. C., Larson, F. C., Tust, R. H. 1954. *Proc. Soc. Exp. Biol. Med.* 86:137–40
9. Becker, D. V., Prudden, J. F. 1959. *Endocrinology* 64:136–48
10. Pitt-Rivers, R., Stanbury, J. B., Rapp, B. 1955. *J. Clin. Endocrinol. Metab.* 15:616–20
11. Lassiter, W. E., Stanbury, J. B. 1958. *J. Clin. Endocrinol. Metab.* 18:903–6
12. Braverman, L. E., Ingbar, S. H., Sterling, K. 1970. *J. Clin. Invest.* 49:855–64
13. Pittman, C. S., Chambers, J. B., Read, V. H. 1971. *J. Clin. Invest.* 50: 1187–96
14. Surks, M. I., Schadlow, A. R., Stock, J. M., Oppenheimer, J. H. 1973. *J. Clin. Invest.* 52:805–11
15. Braverman, L. E. et al 1973. *J. Clin. Invest.* 52:1010–17
16. Chopra, I. J. 1974. *Ann. Intern. Med.* 81:73–75
17. Nicoloff, J. T., Low, J. C., Dussault, J. H., Fisher, D. A. 1972. *J. Clin. Invest.* 51:473–83.
18. Chopra, I. J., Fisher, D. A., Solomon, D. H., Beall, G. N. 1973. *J. Clin. Endocrinol. Metab.* 36:311–16
19. Stock, J. M., Surks, M. I., Oppenheimer, J. H. 1974. *N. Engl. J. Med.* 290:529–33
20. Escobar del Rey, F., Morreale de Escobar, G. 1961. *Endocrinology* 69: 456–65
21. Braverman, L. E., Ingbar, S. H. 1962. *Endocrinology* 71:701–12
22. Hershman, J. M. 1964. *J. Clin. Endocrinol. Metab.* 24:173–79
23. Furth, E. D., Rives, K., Becker, D. V. 1966. *J. Clin. Endocrinol. Metab.* 26:239–46
24. Nicoloff, J. T. 1970. *J. Clin. Invest.* 49:267–73
25. Oppenheimer, J. H., Schwartz, H. L., Surks, M. I. 1972. *J. Clin. Invest.* 51:2493–97
26. Geffner, D. L., Azukizawa, M., Hershman, J. M. 1974. *Clin. Res.* 22:339A (Abstr.)
27. Saberi, M., Sterling, F., Utiger, R. 1974. *Program Fifty Sixth Ann. Meet. Endocrine Soc.*, p. A-90 (Abstr.)
28. Frumess, R. D., Larsen, P. R. 1974. *Program Fifty Sixth Ann. Meet. Endocrine Soc.*, p. A-91 (Abstr.)
29. Schwartz, H. L., Surks, M. I., Oppenheimer, J. H. 1971. *J. Clin. Invest.* 50:1124–30
30. Chopra, I. J., Solomon, D. H., Chua Teco, G. N. 1973. *J. Clin. Endocrinol. Metab.* 36:1050–57
31. Evered, D. C., Ormston, B. J., Smith, P. A., Hall, R., Bird, T. 1973. *Brit. Med. J.* 1:657–62
32. DeLange, F., Camus, M., Ermans, A. M. 1972. *J. Clin. Endocrinol. Metab.* 34:891–95
33. Kochupillai, N. et al 1973. *Lancet* 1:1021–24
34. Portnay, G. I. et al 1974. *J. Clin. Endocrinol. Metab.* 39:191–94
35. Sullivan, P. R. C., Bollinger, J. A., Reichlin, S. 1973. *J. Clin. Invest.* 52:83a (Abstr.)
36. Nomura, S., Pittman, C. S. 1974. *J. Clin. Invest.* 53:57a (Abstr.)
37. Chopra, I. J., Solomon, D. H., Chopra, U., Young, R. T., Chau Teco, G. N. 1974. *Clin. Res.* 22:337A (Abstr.)
38. Bermudez, F., Surks, M. I., Oppenheimer, J. H. 1974. *Clin. Res.* 22:335A (Abstr.)
39. Chopra, I. J. 1974. *J. Clin. Invest.* 53:16a (Abstr.)
40. Lieblich, J. M., Utiger, R. D. 1973. *J. Pediat.* 82:290–92
41. Fisher, D. A., Dussault, J. H., Hobel, C. J., Lam, R. 1973. *J. Clin. Endocrinol. Metab.* 36:397–400
42. Abuid, J., Stinson, D. A., Larsen, P. R. 1973. *J. Clin. Invest.* 52:1195–99
43. Pittman, C. S., Pittman, J. A. 1974. *Handb. Physiol.* 3:233–53

Copyright 1975. All rights reserved

THE TOTAL CORRECTION OF CONGENITAL HEART DISEASE IN INFANTS

❖7165

Eoin Aberdeen, M.B., F.R.C.S.
Division of Cardiothoracic Surgery, Department of Surgery, Mount Sinai Medical Center, New York, NY 10029

THE NATURAL AND UNNATURAL HISTORY OF CONGENITAL HEART DISEASE

The successful correction of congenital heart disease by intracardiac operations performed during the first year of life has been reported in recent years from a number of centers around the world (1–13). Total correction can, in many instances, lead to the expectation of normal heart function for a normal life span.

The factors which influence survival of infants and children with congenital heart disease have not been completely recognized, but it is clear that the more severe the anomaly the younger the age at which heart failure or death may occur. Age in congenital heart disease is inversely related to severity. Simply to have stayed alive indicates that an infant has passed a basic test of physiologic functions. Not to survive is to fail the "trial of living," and the younger the age at which survival is threatened, the more likely it is that the lesion is complex. Thus the incidence of different types of congenital heart disease will vary with the age at which the patient is examined. By the age of 6 months many complex or severe lesions have been filtered out by death, so the incidence of types of congenital lesions presenting at one year of age is very different from the incidence noted at birth [much of the data demonstrating the filtering effect of age are assembled in a masterly review by Hoffman (15)]. Diagnosis has been steadily improving over the last two decades, so recent studies on the incidence of different types of congenital heart disease (CHD) from centers having a special interest and expertise in the subject, such as the New England Regional Infant Cardiac Program (16), are more likely to be valid than the studies reported a decade or more ago. The incidence of the eleven most common forms of CHD observed over a four-year period in six New England states is shown in Table 1.

Table 1 New England regional infant cardiac program: Diagnostic distribution of top eleven lesions

Lesion	Number	Percentage	
1. Ventricular septal defect	233	14.9	50%
2. Transposition of the great arteries	167	10.8	
3. Tetralogy of Fallot	137	8.7	
4. Hypoplastic left heart	122	7.8	
5. Patent ductus arteriosus	116	7.4	
6. Coarctation of the aorta	111	7.1	25%
7. Endocardial cushion defect	75	4.8	
8. Malposition	58	3.7	
9. Single ventricle	49	3.1	
10. Pulmonary atresia	48	3.1	
11. Pulmonary stenosis	48	3.1	
No significant heart disease (mostly pulmonary problems)	38	2.7	25%
Other	362	22.8	
TOTAL	1,564[a]	100.0	

[a] Missing six patients.

Examples of Natural Selection Affecting Surgical Results

Some lesions have a high risk because they occur in combination with one or more other congenital lesions, either of the heart or other organs. The most obvious examples of the combination of cardiovascular lesions are the combination of extra- and intracardiac lesions which occur with coarctation of the aorta and/or patent ductus arteriosus.

COARCTATION OF THE AORTA In a series reported from the Hospital for Sick Children, Great Ormond Street, London, of 84 infants who required operation for correction of coarctation of the aorta before the age of 6 weeks, 80 had additional cardiovascular lesions. Of a total of 88 infants under the age of 12 months who died at the time of, or following, operative correction of the coarctation, 87 had additional cardiovascular lesions. Overall experience showed that in cases with coarctation of the aorta it was rare that coarctation alone required operation early in life, and, if it did, the patient almost always survived (14, 17–19).

PATENT DUCTUS ARTERIOSUS This abnormality reveals a natural history profile which is very similar to that of coarctation (20). Of 158 infants requiring duct closure before the age of 6 months, 51% (81 cases) had additional intracardiac lesions, compared to 691 cases over one year of age of whom only 13% (88 cases) had additional intracardiac lesions. Again, death did not occur in the older, uncomplicated cases requiring treatment, but only in cases with additional severe congenital deformities of the heart or other organs, or among the older children, in those with severe pulmonary hypertension.

VENTRICULAR SEPTAL DEFECT This lesion provides another interesting example of natural history affecting surgical results ("unnatural history"). Of a series of 146 cases which had pulmonary artery (PA) constriction (banding) at the Hospital for Sick Children, London, in the years 1957–1966, about two thirds had intracardiac lesions additional to ventricular septal defect (VSD) (21). Of 22 cases operated before the age of 6 weeks, only 2 had VSD without other intracardiac lesions, whereas of 49 operated on between the ages of 3 and 6 months, 20 had VSD without other intracardiac lesions. A sharp contrast in survival rates was apparent in this early series, the hospital mortality rate for banding of cases with VSD alone was 2% (one death in 44 cases), but the mortality rate for cases with VSD and an additional lesion such as a large patent ductus, coarctation of the aorta, or atrial septal defect (ASD) was between 54% and 77%—an astonishing contrast. The low mortality for treating VSD alone was pleasantly surprising because some of the cases had been so ill before operation. During the next two years the mortality rate had increased to 7% for treating VSD without other defects by PA banding (22), and it was recognized that for cases with isolated single VSD, primary correction is preferable, but such cases rarely required correction before the age of 3 months.

A further point of interest concerning the natural history of this series of cases with VSD occurred when the corrective operation was performed (22). Of 22 cases corrected who had no lesion other than VSD, 8 had multiple VSDs (either all in the muscular septum, or a membranous septal defect combined with muscular defects) and 12 had membranous defects of 20-mm diameter or greater (such as 40 X 20 mm or 35 X 25 mm). This was a much higher incidence of very large VSDs and multiple VSDs than had been encountered in previous experience of VSD repair, and strongly suggested that the operation of pulmonary artery banding had interfered with the natural selection and had salvaged cases with very large and/or multiple defects who otherwise would have died. In this series of second-stage VSD corrections, no case with a moderate-sized VSD, single or multiple, died at correction. However, the death rate was high (7 of 18 cases) in those with very large membranous defects. (Postoperative dysrhythmia was a frequent problem then, but one which should be much less of a problem now that a bundle-detector probe is available at operation.)

So, with vetricular septal defect, as with patent ductus and with coarctation, there would appear to be a considerably wider spectrum of the natural history than has sometimes been appreciated, and the younger the age of presentation, the more likely it is that the problem is complex.

FALLOT'S ANOMALY Fallot's anomaly (tetralogy) is another congenital heart lesion where, in the past, the risk of palliation has proved higher in the early weeks and months of life (17). This, however, was not associated with a high incidence of additional other anomalies in the very young, but was associated with the degree of pulmonary arterial, pulmonary valve, and infundibular obstruction, and with the technical problems of small vessel anastomosis. In this instance it was the infant with the hypoplastic infundibulum, pulmonary valve ring, or pulmonary arteries which usually presented at an earlier age.

In addition, a review of the postmortem specimens of the hearts of children dying in early infancy from Fallot's anomaly or coarctation of the aorta showed a high incidence of ischemic injury to the myocardium, with replacement of the hypertrophied ventricular myocardium by fibrous tissue in 35% of those infants dying with Fallot's anomaly (23). Of those infants dying with coarctation of the aorta, 38% showed these severe changes in the myocardium. Here is a further example of the natural history being affected by the severity of the lesion and the development of a complication (fibrotic replacement of the hypertrophied myocardium).

TOTAL ANOMALOUS PULMONARY VENOUS DRAINAGE This severe circulatory derangement provides yet another example of how the natural history changes the clinical presentation, so that age may be an important indicator of risk. Again quoting the early experience at the Hospital for Sick Children in London, of those infants requiring operation in the first year of life, 18 with drainage of the anomalous pulmonary vein to the left innominate vein had been corrected between the years 1963–1968 (17). Eight of these had been operated on before the age of 3 months and only 1 survived, but of the 10 corrected using cardiopulmonary bypass, and who were between the ages of 3 and 12 months at the time of operation, only 1 died. Why such a sharp contrast? Is this the effect of age, or is some selective factor at work? The selective factor was well defined by Behrendt et al (24), in a study which included the same case material. They found that in cases draining to the left innominate vein, survival was very closely related to the degree of pulmonary hypertension. If the right ventricle (RV) pressure was greater than systemic, survival was very unlikely, whereas if the RV pressure was only 75% of systemic pressure or less, survival was very likely. This study also demonstrated that severe degrees of obstruction to the anomalous vein, causing RV pressure to be greater than systemic, were almost always the result of obstruction in the ascending anomalous vein, either intrinsic or extrinsic. So, the explanation of the contrast in results of this early experience was that those with obstruction to the anomalous pulmonary vein had very severe pulmonary hypertension and developed severe heart failure *before* 3 months of age, and were the only cases that did so, whereas those who presented later were unobstructed, had much less pulmonary hypertension, and were good risk cases for operation. Age at operation did separate the groups into the high and low risk, but age was not the cause of the risk; rather it was pulmonary vein obstruction and severe pulmonary hypertension.

A careful study of this problem by Parr et al (25), in which cardiac output was studied, suggested that a few infants with total anomalous pulmonary venous drainage (TAPVD) have a small-capacity left atrium which results in low cardiac output. Data concerning pulmonary hypertension were not included in this report, but it could be that those with severe pulmonary hypertension are the same infants in whom compliance of the left side of the heart is a problem.

TRANSPOSITION OF GREAT ARTERIES This very severe anomaly has proved in the past, if untreated, to be as lethal as TAPVD—that is, about 85% of infants born with either of these anomalies will be dead within the first year of life. The natural history of transposed great arteries (TGA) is very much affected by the presence of other anomalies (26, 27), but in the opposite direction to lesions already discussed. Transposition, *without* any other abnormalities, is incompatible with survival for more than a few days, unless some other anomaly is present which allows some mixing of the pulmonary and systemic circulations. An ASD, or at least a patent foramen ovale, is the commonest defect, and short term survival is improved by a VSD (27). However, the best survival without surgery occurs in TGA with pulmonary stenosis and VSD and ASD.

Transposition of the great arteries has proved to be one of the conditions most suited to correction in the first year of life, and this will be discussed further under results of treatment.

PULMONARY VASCULAR OBSTRUCTIVE DISEASE The most feared long-term complication of congenital heart disease is the development of severe pulmonary vascular obstructive disease. Factors which are known to contribute to an increase in pulmonary vascular resistance are an increase in the pulmonary pressure either of the pulmonary arteries or the pulmonary veins and arteries, an increase in the pulmonary blood flow, an increase in the pulmonary artery oxygenation, systemic arterial desaturation, or an increase in the hematocrit. All these factors are to be found in transposition of the great arteries with a large VSD, and this is the form of congenital heart disease most likely to develop pulmonary vascular disease. Few infants with this combination of lesions reach their first birthday without developing a severe degree of pulmonary vascular disease.

A large VSD with normal great arteries may also develop severe pulmonary vascular disease, but not as rapidly as in cases with transposed great arteries, and the onset is unusual before the age of 2 years and rare before the age of 1 year.

An Example of the Effect of Selection on the Results of Open Heart Surgery in Infants

A comparison of the results of treatment of different groups of cases may have little meaning unless the cases are comparable in their natural history. A good example of this may be observed in a review of the first 70 infants treated by open heart surgery at the Hospital for Sick Children in London (treated between 1963 and 1970). Of the first 15 cases treated, only 3 survived, and of the last 15 cases, 14 survived. At first glance this seems obvious evidence of a great improvement. However, although valuable experience and expertise were acquired during this period, there were few changes in such important details as operative methods, intensive care technique, and perfusion technique. If innate risks of the cases being treated are examined, it is apparent that the difference in results is explained by the natural history of the cases being treated. At first only desperately ill infants were chosen for corrective operation.

The first 15 cases included 14 cases with TAPVD of which:

9 had severe pulmonary hypertension,
1 had multiple other anomalies,
1 had anuria and severe respiratory failure, and
1 had congenital mitral stenosis and endocardial-fibro-elastosis.

The last 15 cases were operated on in 1970, at a time when it had become apparent that cardiopulmonary bypass could be used for sick infants with a very low innate mortality risk from the bypass itself. These last 15 cases happened to be a group of nearly ideal subjects for infant open heart surgery and included:

9 infants with simple TGA (plus ASD) aged 3–11 months,
4 infants with TAPVD (3 had only moderate pulmonary hypertension),
1 with pulmonary valve stenosis, and
1 with aortic valve stenosis.

Had the two series been reversed in time, it is likely that the results would have been little changed, and then it would seem that the performance of the entire team had deteriorated badly. In fact, it is simply natural selection, or the natural history, affecting the surgical results. For these reasons many of the reports of the results of open heart surgery for series of infants may need to be analyzed in detail before meaningful conclusions can be made.

A method of allowing a meaningful assessment of the results of open heart operations in infants is to compare the survival rates of some well-defined lesions in which survival rates of better than 90% have been recorded by a number of centers. Such good risk cases include complete TGA with intact ventricular septum, single ventricular septal defects (not a partial absence of septum or single ventricle), TAPVD without severe pulmonary hypertension, pulmonary valve stenosis, aortic valve stenosis (not in the first 10 days of life), and perhaps Fallot's anomaly. The crucial test for a good unit is that avoidable deaths should not occur, and an estimate of survival in good risk cases would serve as an indicator of the performance of a unit, just as in general pediatric surgical centers, the survival rate of infants with esophageal atresia serves as an index of performance (good risk infants have had a survival rate of 95% for more than the last two decades in some units). What is accepted as an unavoidable death is becoming less with increasing experience, and perhaps the advent of an effective form of cardiac assistance in infants in the first few days after operation will increase the survival rate still further in complex lesions.

TECHNIQUES AND THE DEVELOPMENT OF OPEN HEART SURGERY IN INFANTS

Successful open heart surgery in infants under one year of age has a relatively long history. Two centers were first to show that open heart surgery using cardiopulmonary bypass could be performed safely in infants; these were the Mayo Clinic (2) and Ann Arbor (1). Kirklin & DuShane (2) had achieved the survival of 20 of 21 cases, aged 6–24 months, having repair of VSD in the years 1959–1960. Sloan et al (1) reported in 1962 the results of operation on 56 patients less than 2 years old who had open heart surgery at Ann Arbor from 1957–1961. Of these patients, 30 had VSD closure, and the only deaths were those under 4 kg body weight or those with severe pulmonary vascular disease. There were no deaths among the other infants having VSD repair.

Included in the remarkable report of Sloan et al (1) was the first successful correction of total anomalous pulmonary venous drainage to below the diaphragm (to the portal vein). On June 9, 1961, the baby was surface cooled to 29°C. The circulation was arrested for 24 min while the anastomosis was made and the infant then rewarmed using cardiopulmonary bypass. This identical method of using hypothermia was reported from Japan in 1967, and has become known as the Kyoto technique (5). It was later applied by Barratt-Boyes (8) to produce a series of brilliant results of total correction in the first weeks and months of life, which served as a great stimulus to other centers to attempt total correction at an early age.

Very few centers have published results to equal those from Auckland, but several have recorded results which were an improvement on their own previous experience using cardiopulmonary bypass. However, it is clear that this technique of hypothermic arrest does not necessarily lead to excellent results in poor risk cases (9, 13). Much of the data concerning total correction in infants has been reported in insufficient detail to assess the effect of the natural history. The present position is that excellent results have been obtained both using cardiopulmonary bypass with or without hypothermia, and using hypothermic arrest supported by cardiopulmonary bypass. The technique initially developed by Horiuchi et al (4) of using surface cooling and surface rewarming with cardiac arrest, but not using cardiopulmonary bypass, has had a more limited application, and has not produced results of the quality of the methods already mentioned. Another use of hypothermia which employs two mechanical pumps, one for the pulmonary and one for the systemic circulation, so that the lungs can be used for ventilation while achieving a deep hypothermic state for cardiac arrest (28), has proved helpful but not as effective in infants as other methods.

The techniques which involve prolonged periods of circulatory arrest have posed the question of whether minor degrees of brain damage may occur, which will not become measurable until several years after operation. Postoperative pulmonary complications may be less after hypothermic arrest techniques. Renal failure may be a problem, however, and Subramanian (29) reported 3 (9%) cases dying of renal failure in 35 infants undergoing correction of TGA. Some infants may be so ill before operation that they require artificial ventilation and even catecholamine support of the circulation. For these patients, postoperative care is usually as important as the operative management.

RESULTS OF TREATMENT

Some conditions can only be managed by corrective surgery. These are:

1. Severe aortic valve stenosis, severe pulmonary valve stenosis, or atresia with intact ventricular septum.
2. Patent ductus arteriosus and/or coarctation of the aorta.
3. Cardiac tumors.
4. Atrial septal defect (rarely).
5. Anomalous left coronary artery.

Some conditions do not permit correction at any age, and, for these cases (which include severe left heart hypoplasia, mitral or aortic valve atresia, some forms of tricuspid atresia with or without pulmonary atresia, severer forms of Ebstein's anomaly, and some forms of single ventricle), palliative surgery is the only help available.

Several conditions present with the possibility of having either palliative or corrective surgery and attention will be directed at the results which have been achieved in the treatment of these lesions.

Ventricular Septal Defect

The results achieved at the Mayo Clinic by Kirklin and at Ann Arbor by Sloan in the late 1950s have already been referred to, and have hardly been improved on by the best of recent results. Closure of a ventricular septal defect may be life-saving for a child in severe heart failure, but there are also long term concerns which favor the early closure of ventricular septal defect. The development of pulmonary vascular disease in association with ventricular septal defect has been well studied. After closure of the ventricular septal defect, between 5 and 10% of those children who have moderate to severe pulmonary hypertension before operation will continue to increase their pulmonary vascular resistance after operation (30–32). A large VSD may result in left ventricular hypertrophy and depression of the contractile state as a result of long-standing volume overload, which may not return to normal after VSD closure (33, 34). Studies of the pulmonary arterial tree in infants and children with ventricular septal defect show that the hypertrophy of the smooth muscle is apparent from the age of 2 months onward and intimal proliferation becomes apparent between 6 and 12 months after birth (35–37).

Does early closure of VSD prevent the development of severe pulmonary vascular disease? Studies by DuShane (38) at the Mayo Clinic suggest that closure before the age of 2 years does prevent the development of severe pulmonary vascular disease. A careful review of the Toronto experience by Keith et al (39) shows that the only cases which will assuredly not develop increasing pulmonary vascular resistance are those in whom the pulmonary systemic flow ratio is less than 2:1 and in whom the pulmonary vascular resistance is normal. Even those cases with a normal vascular resistance but in whom the flow ratio was greater than 2:1 developed a severe increase in vascular resistance in 4 cases of 64. Once severe pulmonary vascular disease has developed, children without operation may live longer than those who have successfully survived VSD closure (40).

Spontaneous closure of ventricular septal defects was observed in 21% of the cases studied by Keith et al (39), and in a smaller series studied by Hoffman and Rudolph, 52% either closed or became smaller (41). Despite the large number of ventricular septal defects which closed spontaneously, the threat of developing severe pulmonary vascular obstructive disease is major and threatens 5 to 10% of patients with moderate-sized or larger ventricular septal defects. Since the risk of correcting these defects should be less than 5%, operation should be indicated in any patient in whom the pulmonary systemic flow ratio is greater than 2:1 and operation should certainly be done within the first 2 years and probably by preference within the first 12 months of life, provided that the survival rate in the center concerned is better than 95%.

Banding of the pulmonary artery may still have a limited place for infants in the first months of life with a common ventricle, a very large ventricular septal defect, large multiple ventricular septal defects, or with an atrio-ventricular canal. A precise anatomical diagnosis is clearly required before a decision concerning operation is made.

Transposition of Great Arteries

This condition, like TAPVD, is one of the most lethal forms of correctable congenital heart disease—without correction 85% of infants will die in their first year. The introduction of a widely applicable operation by Mustard in 1964 (42) was a milestone. The next important contribution was the introduction of balloon atrial septostomy by Rashkind & Miller in 1966 (43), because this allowed an atrial septal defect to be created even in the newborn, at a very low risk. The surgical creation of an atrial septal defect using the Blalock-Hanlon technique (44) has been shown to produce a survival rate of over 95% in infants with TGA over the age of 3 months (45, 46). However, observation of large groups of infants with TGA following balloon atrial septostomy show a continuing mortality rate of 10–15% each year for the first 2–3 years (47). Vlad & Lambert (48) observed satisfactory long term palliation following balloon septostomy in only one half of their cases, when allowing for complications such as cerebrovascular accident and the development of pulmonary vascular obstructive disease.

The possible value of early correction of transposition is apparent from these observations. For these reasons the policy of correction in the first year of life was begun at the Hospital for Sick Children in London in 1967 (49). By 1970, of 17 cases having elective operation in the first year of life, 16 survived, and 1 of 2 cases operated upon as emergencies survived (7). This series has been extended since then by Stark et al (50) and there have been no deaths in the next 43 infants operated on between the ages of 3 weeks and 11 months. This mortality rate of 4% compares with the mortality rate in the same hospital of 10% for 143 cases of simple transposition operated on over the age of 1 year. A few of the smallest infants were operated on using hypothermic arrest and the remainder using cardiopulmonary bypass with some hypothermia. A high survival rate for Mustard operation in infants has been recorded by Barratt-Boyes et al (8), Kilman et al (51), Subramanian (29), and Nadas et al (16). It is of great interest that not only can very good survival rates be achieved using several different techniques of cardiac support, but that the survival rate seems to be higher in infants than in older children. Follow-up studies up to 7 years after operation indicate that the children are growing satisfactorily and have not developed late obstruction of caval or pulmonary vein flow paths when pericardium was used for the interatrial division. Experience using Dacron for the atrial dividing patch did produce a large number of complications as a result of a severe fibrosis (50), and pericardium is now preferred by most for the Mustard operation in the first year of life.

Complex forms of TGA have not yielded such encouraging results as yet, and palliative surgery still has a place in their management.

Fallot's Anomaly

Whether it is best to correct Fallot's anomaly in the first year of life or to use a palliative operation is not yet clear. Good results in the correction of Fallot's in the first year of life have been reported by Baffes (6), Starr et al (52), Barratt-Boyes (53), Pacifico (54), and Nadas (16). It would seem that total correction of Fallot's anomaly can be achieved in infants under the age of a year with a survival rate of better than 90%. However, a large number of these infants have had an infundibular outflow patch inserted and only time will tell whether this is of significance later.

The risks of systemic-to-pulmonary artery shunt operations have also steadily diminished in recent years. The Blalock-Taussig operation has been done with mortality rates of about 2% in children over a year of age for some years now. The risks in younger children have seemed higher, but recent experience using microsurgical techniques and 7–0 Prolene sutures have shown that even in the first days of life a reliable shunt can be constructed at a very low risk. In the writer's personal experience, 17 infants with Fallot's anomaly or pulmonary valve atresia have had a shunt performed in the first 6 days of life in the last 3½ yr, with 15 surviving and no late shunt closures so far. Mobilizing the right common carotid artery into the neck allows the right subclavian to be drawn down and so permits a 3-mm anastomosis in most infants. Thus the writer now prefers the Blalock-Taussig shunt to other types of shunt should a shunt be required at any age. Good results for Blalock-Taussig shunts in infants have also been recorded by others with mortality rates between 0 and 4% (55–57).

Pulmonary atresia with intact ventricular septum, presenting in the first few days of life, has so far yielded best results when a balloon septostomy has been followed by a systemic-to-pulmonary artery shunt (58, 59) and this may be followed by pulmonary valvotomy within minutes, days, or weeks.

Total Anomalous Pulmonary Venous Drainage

The mortality rate for this condition is closely associated with the degree of pulmonary hypertension. Gathman & Nadas (60) noted that only 2 of 25 patients with pulmonary hypertension survived through the first year, whereas 19 survived of 22 without pulmonary hypertension. The influence of severe degrees of pulmonary hypertension on the results of surgical correction have already been referred to (24).

Two questions have not been finally answered. One is the place of balloon atrial septostomy in the early management of TAPVD. Some cases have been reported with a fall in pressure from systemic levels to near normal pressures following balloon atrial septostomy (61, 62). Cases with near normal pulmonary pressure have little risk either at operation or by being observed through the first year of life, whereas cases with pulmonary hypertension face high risks. Those with anatomical obstruction to the pulmonary venous return can only be managed by surgical correction, and even then the survival rate may not be high. Perhaps there is an intermediate group in which the pulmonary hypertension is at systemic level, and is not caused by venous obstruction but is possibly the result of a constrictive atrial septal defect or patent foramen ovale, in which balloon septostomy may be the treatment of choice.

The other question concerns the best technique of cardiac support during operation. Some excellent results have been achieved using hypothermia and circulatory arrest, although the results have varied considerably among different centers, but insufficient data are given to determine whether these results are due to differences in surgical technique or to the natural history of the disease (63, 64). Excellent results have been recorded using cardiopulmonary bypass also (65). The technical advantages obtained by circulatory arrest are most useful when anastomosing the anomalous vein to the left atrium.

The form of TAPVD With the highest mortality rate has been when the anomalous vein drained below the diaphragm. Results of recent treatment with hypothermia and circulatory arrest suggest that the technique may have its most important application in this group.

SUMMARY

If radical correction of congenital heart disease can be performed with a low risk in the first year of life the advantages are obvious. Total correction avoids the latent risks of the underlying lesion and the risks of palliative surgery, and relieves the parents and family of the psychological pressure of a major illness. Palliative surgery, although offering survival, may include the long term problems of the palliative operation itself, and the possibility that cases may be lost to further treatment or follow-up examination after a successful palliative operation. For these reasons, to be desirable, a palliative operation should offer a survival chance that is at least 10% better than the corrective procedure in that patient at that time. That is, any corrective procedure with a mortality rate less than 10% is to be preferred at any age to a palliative operation. An accurate assessment of risk demands a complete diagnostic study, and a knowledge of the natural history of that form of congenital heart disease.

The problems of infants with congenital heart disease are not primarily caused by age or size. The main problem is that of natural selection.

Literature Cited

1. Sloan, H., McKenzie, J., Morris, J. D., Sterne, A., Sigmann, J. 1962. Open heart surgery in infancy. *J. Thorac. Cardiov. Surg.* 44:459
2. Kirklin, J. H., DuShane, J. H. 1961. Repair of ventricular septal defect in infancy. *Pediatrics* 27:961
3. Ashmore, P. G. 1963. Cardiopulmonary bypass in the infant. *Surgery* 54:361
4. Horiuchi, T. et al 1963. Radical operation for ventricular septal defect in infancy. *J. Thorac. Cardiov. Surg.* 46:180
5. Hikasa, Y. et al 1967. Open heart surgery in infants with an aid of hypothermic anesthesia. *Arch. Jap. Chir.* 36:495
6. Baffes, T. G. 1968. Total body perfusion in infants and small children for open heart surgery. *J. Pediat. Surg.* 3:551
7. Aberdeen, E. 1971. Correction of uncomplicated cases of transposition of the great arteries. *Brit. Heart J.* 33: 66–68
8. Barratt-Boyes, B. G., Simpson, M., Neutze, J. M. 1971. Intracardiac surgery in neonates and infants using deep hypothermia with surface cooling and limited cardiopulmonary bypass. *Circulation* 43 and 44 (Suppl. I):25
9. Subramanian, S., Wagner, H., Vlad, P., Lambert, E. 1971. Surface induced deep hypothermia in cardiac surgery. *J. Pediat. Surg.* 6:612
10. Malm, J. R. et al 1972. Techniques in management of infants requiring cardi-

opulmonary bypass. *Birth Defects: Original Article Series, Vol. VIII, No. 1*

11. Bonchek, L. J., Anderson, R. P., Wood, J. A., Chapman, R. D., Starr, A. 1974. Intracardiac surgery with extracorporeal circulation in infants. *Ann. Thorac. Surg.* 17:280
12. Breckenridge, I. M. et al 1973. Open-heart surgery in the first year of life. *J. Thorac. Cardiov. Surg.* 65:58
13. Rittenhouse, E. A., Hitoshi, M., Dillard, D. H., Merendino, K. A. 1974. Deep hypothermia in cardiovascular surgery. *Ann. Thorac. Surg.* 17:63–98
14. Aberdeen, E. 1968. Cardiac surgery in the first months of life. In *Modern Trends in Cardiac Surgery,* 155–81. London: Butterworth
15. Hoffman, J. I. E. 1968. Natural history of congenital heart disease: problems in its assessment with special reference to ventricular septal defects. *Circulation* 37:97–124
16. Nadas, A. S., Fyler, D. C., Castaneda, A. R. 1973. The critically ill infant with congenital heart disease. *Mod. Conc. Cardiovasc. Dis.* XLII: 53–58
17. Stark, J., Hucin, B., Aberdeen, E., Waterston, D. J. 1971. Cardiac surgery in the first year of life: experience with 1,049 operations. *Surgery* 69:483–97
18. Easthope, R. N., Tawes, R. L. Jr., Bonham-Carter, R. E., Aberdeen, E., Waterston, D. J. 1969. Congenital mitral valve disease associated with coarctation of the aorta: a report of 39 cases. *Am. Heart J.* 77:743–54
19. Tawes, R. L. Jr., Aberdeen, E., Waterston, D. J., Bonham-Carter, R. E. 1969. Coarctation of the aorta in infants and children: a review of 333 operative cases, including 179 infants. *Circulation* XXXIX and XL (Suppl. I): 173–84
20. Panagopoulos, P. G., Tatooles, C. J., Aberdeen, E., Waterston, D. J., Bonham-Carter, R. E. 1971. Patent ductus arteriosus in infants and children. *Thorax* 26:137–44
21. Stark, J., Aberdeen, E., Waterston, D. J., Bonham-Carter, R. E., Tynan, M. 1969. Pulmonary artery constriction (banding): a report of 146 cases. *Surgery* 65:808–18
22. Stark, J. et al 1970. Repair of intracardiac defects after previous constriction (banding) of the pulmonary artery. *Surgery* 67:536–47
23. Tawes, R. L. Jr., Berry, C., Aberdeen, E., Graham, G. R. 1969. Myocardial ischemia in infants: its role in three common congenital cardiac anomalies. *Ann. Thorac. Surg.* 8:383–90
24. Behrendt, D. M., Aberdeen, E., Waterston, D. J., Bonham-Carter, R. E. 1972. Total anomalous pulmonary venous drainage in infants. I. clinical and hemodyndamic findings, methods, and results of operation in 37 cases. *Circulation* 46:347
25. Parr, G. V. S., Kirklin, J. W., Pacifico, A. D., Blackstone, E. H., Louridsen, P. 1974. Cardiac performance in infants after repair of total anomalous pulmonary venous correction. *Ann. Thorac. Surg.* 17:561–73
26. Hanlon, C. R., Blalock, A. 1948. Complete transposition of the aorta and the pulmonary artery: experimental observations on venous shunts as corrective procedures. *Ann. Surg.* 127:385–94
27. Liebman, J., Cullum, L., Belloc, N. B. 1969. Natural history of transposition of the great arteries: anatomy and birth and death characteristics. *Circulation* XL:237–62
28. Drew, C. E., Anderson, F. M. 1959. Profound hypothermia in cardiac surgery: report of three cases. *Lancet* 1:748
29. Subramanian, S., Wagner, H. 1973. Correction of transposition of the great arteries in infants under surface-induced deep hypothermia. *Ann. Thorac. Surg.* 16:391–401
30. Cartmill, T., DuShane, J. W., McGoon, D., Kirklin, J. W. 1966. Results of repair of ventricular septal defect. *J. Thorac. Cardiov. Surg.* 52:486
31. Birks, W., Reidemeister, C. 1971. Results of surgical treatment of ventricular septal defect. *Brit. Heart J. Suppl.* 33:88
32. Gotsman, M. S., Beck, V. T., Barnard, C. N., Schrire, V. 1969. Haemodynamic studies after repair of ventricular septal defect. *Brit. Heart J.* 31:63
33. Jamarkani, J. M. M., Graham, T. P. J., Canent, R. V. Jr., Capp, N. P. 1971. The effect of corrective surgery on left heart volume and mass in children with ventricular septal defect. *Am. J. Cardiol.* 27:254
34. Jamarkani, J. M., Graham, T. P. Jr., Canent, R. V. Jr. 1972. Left ventricular contractile state in children with successfully corrected ventricular septal defect. *Circulation Suppl. I* 45:102–10
35. Wagenvoort, C. A., Neufeld, H. N., DuShane, J. W., Edwards, J. E. 1961. The pulmonary arterial tree in ventricular septal defects. *Circulation* 23:740
36. Wagenvoort, C. A. 1962. The pulmonary arteries in infants with ventricular septal defect. *Med. Thorac.* 19:354–61

37. Naeye, R. L. 1966. The pulmonary arterial bed in ventricular septal defect: anatomic features in childhood. *Circulation* 34:962
38. DuShane, J. W., Kirklin, J. W. 1973. Late results of the repair of ventricular septal defect on pulmonary vascular disease. *Advances in Cardiovascular Surgery,* ed. J. W. Kirklin, 9. New York: Grune & Stratton
39. Keith, J. D., Rose, V., Collins, G., Kidd, B. S. L. 1971. Ventricular septal defect: incidence, morbidity and mortality in various age groups. *Brit. Heart J. Suppl.* 33:81
40. Friedli, B., Kidd, B. S. L., Mustard, W. T., Keith, J. D. 1974. Ventricular septal defect with increased pulmonary vascular resistance: late results of surgical closure. *Am. J. Cardiol.* 33:403–9
41. Hoffman, J. E., Rudolph, A. M. 1965. The natural history of ventricular septal defects in infancy. *Am. J. Cardiol.* 16:634
42. Mustard, W. T. 1964. Successful two-stage correction of transposition of the great vessels. *Surgery* 55:469
43. Rashkind, W. J., Miller, W. M. 1966. Creation of an atrial defect without thoracotomy: a palliative approach to complete transposition of the great arteries. *J. Am. Med. Assoc.* 196:991
44. Blalock, A., Hanlon, C. R. 1948. Interatrial septal defect: its experimental production under direct vision without interruption of the circulation. *Surg. Gynecol. Obstet.* 87:183
45. Aberdeen, E. 1968. Blalock-Hanlon operation and Rashkind procedure. In *Operative Surgery,* 2:193–99. London: Butterworth. 2nd ed.
46. Deverall, P. B. et al 1969. Palliative surgery in children with transposition of the great arteries. *J. Thorac. Cardiov. Surg.* 58:721–29
47. Tynan, M. 1971. Survival of infants with transposition of great arteries after balloon atrial septostomy. *Lancet* 1:621
48. Vlad, P., Lambert, E. C. 1973. Late results of Rashkind's balloon atrial septostomy in transposition. See Ref. 38, p. 29
49. Aberdeen, E., Carr, I. 1968. Transposition of the great arteries. See Ref. 14, pp. 182–99
50. Stark, J., de Leval, M. R., Waterston, D. J., Graham, G. R., Bonham-Carter, R. E. 1974. Corrective surgery of transposition of the great arteries in the first year of life: results in 63 infants. *J. Thorac. Cardiov. Surg.* 67:673–81
51. Kilman, J. W., Williams, T. E. Jr., Kakos, G. S., Craenen, J., Hosier, D. M. 1973. Surgical correction of the transposition complex in infancy. *J. Thorac. Cardiov. Surg.* 66:387–93
52. Starr, A., Bonchek, L. J., Sunderland, C. O. 1973. Total correction of tetralogy of Fallot in infancy. *J. Thorac. Cardiov. Surg.* 65:45
53. Barratt-Boyes, B. G., Neutze, J. M. 1973. Primary repair of tetralogy of Fallot in infancy using profound hypothermia with circulatory arrest and limited cardiopulmonary bypass: a comparison with conventional two-stage management. *Ann. Surg.* 178:406
54. Pacifico, A. D., Bargeron, L. M., Kirklin, J. W. 1973. Primary total correction of tetralogy of Fallot in children less than four years of age. *Circulation* XLVII:1085–91
55. Selmonosky, C. A., Farhangian, D., Folger, G. M., Ellison, R. G. 1972. Palliative shunting operation in tetralogy of Fallot: effects upon the results of total correction. *Ann. Thorac. Surg.* 14:16
56. Goldblatt, E., Discussion of Kerr, A. R., Barratt-Boyes, B. G. 1973. Surgery of tetralogy of Fallot in infancy: comparison of shunt palliation and primary intracardiac repair. In *Heart Disease in Infancy,* ed. B. G. Barratt-Boyes, J. M. Neutze, E. A. Harris, 205. Edinburgh: Churchill-Livingstone
57. Wood, W. C., McCue, C. M., Lower, R. R. 1973. Blalock-Taussig shunts in the infant. *Ann. Thorac. Surg.* 16:454
58. Rowe, R. D., Vlad, P. 1973. Diagnostic problems in the newborn: origins of mortality in congenital cardiac malformation. *Heart Disease in Infancy,* ed. B. G. Barratt-Boyes, J. M. Neutze, E. A. Harris, p. 6. Edinburgh: Churchill Livingstone
59. Miller, G. A. H. et al 1973. Pulmonary atresia with intact ventricular septum and critical pulmonary stenosis presenting in first month of life: investigation and surgical results. *Brit. Heart J.* 35:9–16
60. Gathman, G. E., Nadas, A. S. 1970. Total anomalous pulmonary venous connection: clinical and physiologic observations of 75 pediatric patients. *Circulation* 42:143
61. Miller, W. W. et al 1967. Total anomalous pulmonary venous return: effective palliation of critically ill infants by balloon septostomy. *Circulation* 35(Suppl. 2):189

62. El-Said, G., Mullins, G. B., McNamara, D. G. 1972. Management of total anomalous pulmonary venous return. *Circulation* 45:1240

63. Breckenridge, I. M., de Leval, M., Stark, J., Waterston, D. J. 1973. Correction of total anomalous pulmonary venous drainage in infancy. *J. Thorac. Cardiov. Surg.* 66:447–53

64. Barratt-Boyes, B. G. 1973. Primary definitive intracardiac operations in infants: total anomalous pulmonary venous connection. See Ref. 38, p. 127

65. Gersony, W. M. et al 1971. Management of total anomalous pulmonary venous drainage in early infancy. *Circulation* 43 and 44(Suppl.I):1–19

Copyright 1975. All rights reserved

A DEFINITION OF CEREBRAL DEATH

◆7166

Benjamin Boshes, M.D., Ph.D.[1]
Department of Neurology, Northwestern University-McGaw Medical Center, Chicago, Illinois 60611

Death is a single phenomenon and the legal definition contained in Black's Law Dictionary, Fourth Edition, 1951, reads:

> The cessation of life, the ceasing to exist; defined by physicians as the total stoppage of the circulation of the blood and the cessation of animal and vital functions consequence thereon such as respiration, pulsation, etc. (This is natural death in contradiction to civil death and also to violent death.)

Now, however, with improved knowledge of resuscitation plus available life-sustaining machinery, it is possible to reverse a cardiac standstill or maintain ventilation in a person totally unresponsive and utterly dependent on the respirator. Should the latter be turned off, within a short time the heart would stop. Is a totally unreceptive, unresponsive patient, completely respirator-dependent, alive? Is such a brain capable of survival now or at a later period? What is the effect on the brain of days to weeks on the respirator? The consensus which includes opinions from the Church (1), legal, and ethical sources (2–4) leans towards designating the moment of death at the time that the brain is considered dead. That state has been variously designated as "irreversible coma, brain death, artificial survival, and coma dépassé." The last is well-chosen because the individual is beyond coma, his state is irreversible, and ahead of him is absolute death (5). The Research Group on Death and Dying of the Institute of Society Ethics and Life Sciences (4) formulated its own definition:

> A person should be considered dead if in the announced opinion of a physician, based on ordinary standards of medical practice, he has experienced an irreversible cessation of spontaneous respiratory and circulatory function. In the event that adequate means of support preclude a determination that these functions have ceased, a person will be considered dead if in the announced opinion of a physician, based on ordinary standards of medical practice, he has experienced an irreversible cessation of spontaneous brain function. Death will have occurred at the time when relevant functions ceased.

---

[1] Supported by Contract No. NIH-NINDS-71-2320.

Perhaps the most widely used definition of irreversible coma in this country at present is based on the report of the Harvard Committee of 1968 (6). This describes a patient who is totally unreceptive, unresponsive, who makes no movements, is completely apneic, has no reflexes, and who shows a flat EEG. If all of these criteria are reexamined for and found unchanged after 24 hr, such a patient may be considered as cerebrally dead.

The need for a scientifically sound definition of brain death became increasingly apparent as hospitals and staffs sought surcease from the purposeless waste of money in maintaining life in a person whose brain had no potential for human existence. Needed beds in intensive care units and personnel were immobilized. Relatives became emotional prisoners in the long vigils, while artificial survival was maintained through a respirator. Finally vital organs like kidneys were destroyed in the wait, and potential transplant recipients were denied a chance for useful life. State legislatures were awaiting data to rewrite the legal definition of death.

To answer this need, in the Summer of 1971 a National Institute of Neurological Diseases and Stroke (NINDS) Collaborative Study of Cerebral Survival (Cerebral Death) was launched, involving nine university medical centers. The cases presented herewith are those from the Northwestern University-McGaw Medical Center and while limited in number, reflect the overall findings of the larger report which awaits official publication clearance. A careful protocol was prepared for patient qualification for entering into the study. Controlled procedures were established to provide examination at specific time intervals, a standardized EEG of optimum quality, and certain optional ancillary studies. Children under one year of age were excluded. The entrance criteria required a subject who was totally cerebrally unresponsive and who had been apneic for at least 15 min. Some were brought in with ventilation already assisted. Men dominated in the study, with two age peaks, the teens and the fifties, with trauma and cerebrovascular disease respective causation.

Of 62 cases studied at the Northwestern Center, 38 not only met the criteria for entry, but had complete records for the follow-through. Table 1 represents the etiology of the cerebral unresponsive state and the apnea. Of these, 7 were of cardiac origin, 16 cerebral, 11 metabolic disorders, with 4 others. Most of the cases were entered less than 2 hr after the ictus (Table 2). A history of drug intoxication was found in 3, denied in 29, and unknown in 6 (Table 3), and the barbiturate level as measured in this Center or at the NINDS laboratory was variable (Table 4). Curiously despite the availability of gas/liquid chromatography and thin layer chromotography in our Center, the number of specimens available for analysis was comparatively small and the levels distributed in an unpredictable fashion. The exigencies about the patient made the drawing of blood for such a sample difficult at times, and in other subjects, particularly those who had been in the hospital for a long time, the available record seemed to preclude the need to test for barbiturates.

Temperature ranges were variable, only 4% below 90° F; hyperthermia was also infrequent. The pupillary signs, traditionally utilized as an indicator of death, turned out to be important. Thirteen percent with so-called "small pupils," 1 to 4 mm, survived, whereas of those with large pupils, 5 to 8 mm, there was only 6% survival. No survivors were seen among subjects with pupils over 8 mm. Where the pupillary reaction was absent at the time of entry, only 4% of patients survived; when present,

Table 1 Etiology of comatose state

Cardiac	Cerebral						Exogenous Intoxication	Metabolic Disorder	Neoplasm	Other	Total
	Trauma	Thrombosis	Subarachnoid Hemorrhage	Embolism	Hemorrhage	CNS Infection					
7	4	6	3	–	2	1	1	11	–	3	38

Table 2 Elapsed time between cerebral insult and apnea

Hours	<1.9	2–3.9	4–7.9	8–15.9	16–47.9	>48	Unknown	Total
	20	2	1	1	9	4	1	38

Table 3 History of drug intoxication

History	Yes	No	Unknown	Total
	3	29	6	38

Table 4 Barbiturate levels

μg/ml	0–9	10–29	30–59	60+	Not Done	Unknown	Other Drug	Total
	2	1	—	—	11	24	—	38

21% survived. When the EEG was isoelectric, pupillary reactions were rarely present, 1.6%, but if there was some evidence of electrical brain (biological) activity, 66%, pupillary reactions were retained.

In subjects who met the criteria for entry, total unresponsivity and apnea, audio-ocular reflexes were rarely present. In contrast, the deep tendon spinal reflexes were found in 45%. Other reflexes which were likely to be depressed or absent were pharyngeal, swallowing, and cough.

The presence or absence of deep, superficial, or pathological spinal reflexes or other cephalic reflexes was no indicator of electrical cerebral silence (ECS) or of survival. A Babinski sign was found in only 19%.

The EEG was of enormous importance. The records were technically good and artifacts could usually be identified. If a patient met the criteria for admission and had maintained a flat first EEG for 30 min, we found no instance of survival, even though such a patient might be maintained for hours or days, depending on the wish of the physician or family. Of the 38 patients, 25 died in the course of the coma, a so-called "cardiac" death; 11 died because after an agreed period the respirator was turned off, a "cerebral" death. Two survivors showed incomplete recovery due to residual brain damage. (Table 5) In the multi-institutional study most of the survivors were those with drug intoxication, treated by ventilatory support and/or dialysis. In most the recovery was complete. However, in those who survived after primary neurological insult, stroke, trauma, infection, or metabolic disorder, the recovery was incomplete.

In the Northwestern series, 2 of 38 survived, 5.3%. Most died within 48 hr, 23 of 36 (or the respirator was turned off). Over half of these succumbed within the first 24 hr, 13 of 23, 57% (Table 6).

Twenty-three autopsies were obtained in the 36 deaths, 64%. Table 7 describes the findings. Seven showed a "respirator brain," 30.4%, a condition found in subjects on assisted ventilation. The hemispheres were swollen and dusky, the cerebellum and brain stem tight, molded into each other, herniating into the foramen magnum caudally and into the incisura rostrally. Such a brain was soft and mushy, the cerebellum, particularly its uncus, fragmented. Tonsillar tissue could be found as low as the cauda equina. Microscopically the brain was necrotic, especially the cerebellar folia, brain stem, cortex. The basal ganglia and diencephalon were comparatively spared. Little inflammatory cellular reaction was seen. While such a brain might be expected to produce ECS, this was not always true and some of these patients shortly before death showed evidence of biological activity. Yet the autopsy revealed the so-called respirator brain. Apparently some information from the intact cells was reaching the cortex.

The so-called vital signs obtained in a general physical examination proved to be of little prognostic worth. A low rectal temperature or a very high one, indicating a decompensated thermal mechanism, poikilothermia, was infrequently seen, and those within the normal temperature ranges died as readily as those at the extremes. A falling blood pressure or difficulty in sustaining pressure, or hypertension was no indicator of imminent or delayed death. Less than one fourth of the patients required vasomotor agents. Apparently the blood pressure stabilized a few hours after the

Table 5 Final disposition

Death		Improved Complete	Improved Incomplete	Unknown	Total
Cardiac	Cerebral[a]				
25	11	—	2	—	38

[a] Cases in which resuscitation was stopped, on the basis of presumed cerebral death.

Table 6 Time of death

Hours	1	1–5.9	6–11.9	12–23.9	24–47	48–71	72–95	96–167	1–1.9 Weeks	2–3.9 Weeks	4 Weeks	Total Decd.	Not Decd.
	—	4	5	4	10	2	5	3	2	1	—	36	2

Table 7 Autopsies (23 of 36)

Lesions	None	Respirator Brain	Brain Stem	Diffuse Cerebral	Focal Cerebral	Other	Total Autopsies
	2	7	7	20	10	6	23

ictus. Pulse pressure and pulse rate similarly were of little prognostic value. The systemic vascular system appears to operate independently in the absence of cerebral control. The spinal cord and its autonomic nervous system seems able to maintain adequate general circulation.

ECS or isoelectric EEG was defined as a state where there was no change in electrical potential of cerebral origin over 2 μV during a period of at least 30 min, when recording at a gain of 2 μV to 1 mm pen deflection, from symmetrically placed electrode pairs, 10 or more cm apart. Because most of these tracings had to be made in an Intensive Care Unit while other activity was going on about the subject, certain artifacts crept in. These came from the ventilatory assist apparatus, 60-cycle frequencies, power lines, muscle potentials, the movements attendant to decerebrate or decorticate posturing, and myoclonic seizures. It was possible to eliminate some of the physiologic activity by the use of succinylcholine chloride carefully supervised. Most of the artifacts became familiar to the EEG team at the home site or to the review group.

In the Northwestern series only one patient regained biological activity for a short while; then the EEG became isoelectric again. If a patient with apnea, unresponsiveness, and ECS was maintained on the ventilator, he usually died within one week. If the EEG showed activity he could live up to four weeks. ECS was associated with high mortality. If the criteria had been met and the brain was isoelectric at the outset, 43% of subjects were dead within 24 hr. If electrical activity persisted at the outset and ECS developed later, 51% were dead within the first 24 hr after electrical silence.

The known drug intoxications detected by history, bottles, or identification in the blood and urine produced the most survivors in the large series (not the Northwestern Center). Only two such patients in the large study, with barbiturate intoxication who entered in coma, with areflexia and apnea, showed electrocerebral silence and recovered. Both had small pupils, 2 to 3 mm.

When all of these data are collated, a presumptive definition of cerebral death may be suggested. First the patient must be in coma, totally unresponsive to any form of verbal, movement, or painful stimuli. He must be apneic, unable to breathe on his own for at least 15 min. This subject demonstrates an absence of cephalic reflexes, except vestibular, particularly no audio-ocular, swallowing, coughing, or pharyngeal responses. Pupils are dilated over 4 mm and the EEG maintains cerebral silence for at least 30 min.

From a statistical standpoint a series of 38 is small, and the upper confidence limit would be low. If these data held for a series of 500 patients, the 95% upper confidence limit would be 1.5%. For such a criterion to reach 1–10,000, almost 30,000 subjects would be required, an impossible task in terms of cost, time, and available subjects and personnel. Therefore, other tests are needed to validate these data and among them is a check for cerebral blood flow. The work of Korein (7, 8) and his group at New York University and Bellevue Hospital Center has thrown considerable light on the value of the radioisotope bolus technique as a test to detect circulatory deficits associated with cerebral death. They injected technetium, $^{99m}TcO_4$, into 80 patients in 142 bedside studies. Presence of bolus in the cerebral radioangiogram was associated with significant blood flow. Evidence concerning the latter would suggest at least 20% of normal perfusion. In contrast if the head bolus was absent for a period over 30 min and the clinical picture was that of unresponsiveness, apnea, and an isoelectric EEG, this would be indicative of cerebral death. Here blood flow may be anywhere from 15 to 0%. Sometimes a so-called intermediate tracing was noted but there was no evidence of increased intracranial circulation. It suggested rather that the brain was not receiving the blood; the shunting was into the extracerebral circulation which apparently produced the intermediate effect. Therefore, such a response became equivalent to a "no bolus" effect.

In several instances the diagnostic study of a patient required concomitant four-vessel cerebral angiography. No blood entered the brain above the siphon. It would appear here, with no bolus effect over the brain and good bolus response over the femoral artery, to verify good systemic blood flow, that no cerebral circulation was present. Such data, added to the clinical and electroencephalographic observations, would raise the prediction level of cerebral death. More records of correlations of cerebral blood flow as measured by the radioisotope techniques against four-vessel angiography are needed to firm up the statement. This could raise the clinical, electrical, bolus radioisotope conclusion to a very high confidence measure.

Currently it would appear that if a patient is in coma, totally unresponsive, apneic for 15 min, shows a flat EEG for at least 30 min, has no cephalic reflexes (except vestibular), has dilated pupils over 4 mm, a temperature above 90° F, and there is no evidence by history, clinical examination, or by chemical tests of drug intoxication, such a person has no likely chance of survival. If he shows evidence of drug intoxication, despite the other criteria for cerebral death, and pupils are small, he has the opportunity for survival, often to a functional level. The same is true for patients with temperatures under 90° F.

The current experience would suggest that the Harvard criteria, including the waiting for 24 hr, are too rigid and under proper conditions a diagnosis of cerebral death can be made earlier, with high predictability. In the not too distant future it is hoped that enough additional data will come from correlated arteriographic, radioisotope and other studies to raise this predictability to a higher confidence level.

Literature Cited

1. Pope Pius XII. 1958. The prolongation of life. *The Pope Speaks.* 4:393–98
2. A report by the task force on death and dying of the institute of society, ethics, and life sciences. 1972. An appraisal. *J. Am. Med. Assoc.* 221:48–53
3. Capron, A. N. 1973. Determining death: Do we need a statute? *Hastings Center Report.*
4. Capron, A. N., Kass, L. R. 1972. A statutory definition of the standards for determining human death. *Univ. Pa. Law Rev.* 121:87–118
5. Fischgold, H. 1970. Electroencephalography and organ transplantation. *Electroencephalogr. Clin. Neurophysiol.* 29: 209
6. Beecher, H. K. et al 1968. A definition of irreversible coma—report of the ad hoc committee of the Harvard Medical School to examine the definition of brain death. *J. Am. Med. Assoc.* 205: 337–40
7. Korein, J. 1973. On cerebral, brain and systemic death. Current concepts of cerebral-vascular disease. *Stroke.* 8: 9–14
8. Korein, J., Braunstein, P., Lieberman, A. 1974. Differential diagnosis of cerebral death: Clinical, EEG and portable cerebral blood flow techniques with pathophysiological correlations. *Electroencephalogr. Clin. Neurophysiol.* 36:220

Copyright 1975. All rights reserved

PATHOPHYSIOLOGY OF HEART BLOCK[1]

◆7167

John F. Moran, M.D. and Rolf M. Gunnar, M.D.
Sections of Clinical Electrophysiology and Cardiology, Department of Medicine, Loyola University Stritch School of Medicine, Maywood, Illinois 60153

INTRODUCTION

Older concepts of atrioventricular (A-V) conduction have been clarified recently by data developed through technical advances. These modifications of our ideas of A-V conduction justified a study to revise nomenclature (1). The most important recent advance for study of heart block in the human was the division of A-V conduction into suprajunctional and subjunctional conduction; i.e. the division of the PR interval into the delay from atrial activation to activation of the His bundle (A-H) and the delay from His activation to ventricular activation (H-V). This advance was made possible when the capability developed to record simply the His bundle electrogram in the human (2, 8, 9).

Over 37 years ago, Yater and his colleagues suggested that bilateral bundle branch block might account for 20% of the cases with complete A-V block (3). Unilateral right or left bundle branch block was seen frequently enough to suggest that bilateral bundle disease must be more common than previously thought. In 1955 Rosenbaum & Lepeschkin could find only seven cases of bilateral bundle branch block in the literature but were able to add two cases of their own (4). Lasser et al and Lopez found that 59 and 79% of patients, respectively, with complete A-V block had antecedent electrocardiographic evidence of bilateral bundle branch block (5, 6). The evidence thus supports the thesis that bilateral bundle branch is the most common etiology of complete heart block. However, these and other authors also pointed out that only 10–14% of patients with bilateral bundle branch block ever develop complete A-V block (5, 7).

Knowledge of the precise anatomical relationships between the cardiac conduction system and its blood supply helps in understanding the pathophysiology of complete heart block.

---

[1] Supported in part by U.S. Public Health Service Grant HL05971.

CORONARY ARTERY—ATRIOVENTRICULAR CONDUCTION TISSUE ANATOMY

Although most of the blood supply to the interventricular septum is by way of the penetrating branches of the left anterior descending coronary artery, the first part of the right bundle branch receives twigs from the ramus septi fibrosi, the ramus septi ventriculorum superior, and the ramus cristae, all branches of the right coronary artery (10) (see Figure 1). In some 90% of human hearts, the artery to the atrioventricular node, the ramus septi fibrosi, is a branch of the right coronary artery (11). Posterior septal penetrating branches from the right coronary artery rarely penetrate more than 15 mm into the septum (11). Therefore, nearly the entire blood supply of the interventricular septum comes from penetrating branches of the left anterior descending artery.

In contrast, the A-V bundle (His) frequently has a dual blood supply (13). Frink & James reported that the His bundle was supplied by the AV nodal artery and the first septal branch of the left anterior descending coronary artery in nine of ten normal human hearts (12). In four of these nine cases, the dual blood supply overlapped in the bundle of His. The proximal right and left bundle branches were supplied similarly; in four cases jointly by the A-V nodal artery and the anterior septal branch, and in five cases exclusively by the anterior septal branch. The His bundle was never solely supplied by the anterior septal branch (12).

The second portion of the right bundle branch receives blood from the ramus septi fibrosi but also receives a special supply from the ramus limbi dextri, a branch of

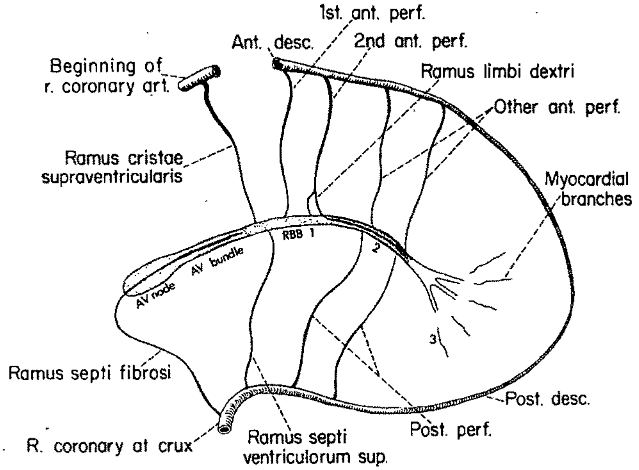


Figure 1 This drawing depicts the relationships of the anterior and posterior descending coronary arteries to right bundle branch, the A-V (His) bundle, and the A-V node in the interventricular septum (adapted from reference 10).

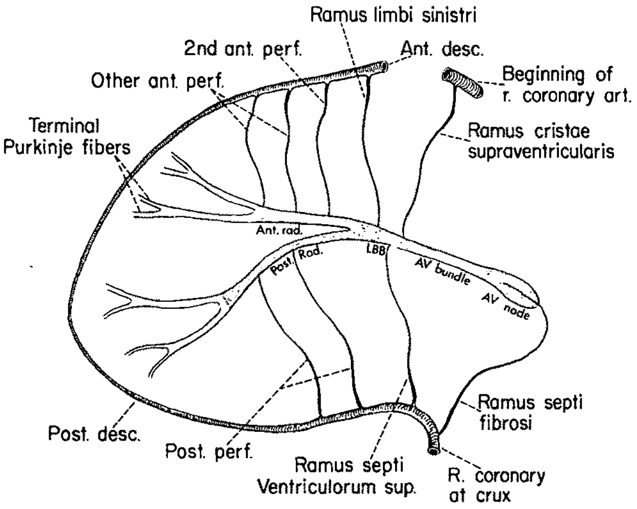


Figure 2 This drawing depicts the relationships of the anterior and posterior descending coronary arteries to the left bundle branch, the A-V (His) bundle, and the A-V node in the interventricular septum (adapted from reference 10).

the second perforating branch of the left anterior descending artery (13) (see Figure 1). The third part of the right bundle branch is supplied by blood vessels to the adjacent myocardium.

The first part of the left bundle branch was supplied entirely by the ramus septi fibrosi in five cases, but received a dual blood supply in four cases (12). It was supplied by the anterior septal branch in only one case in the Frink & James report (12). The ramus septi ventriculorum superior and the ramus cristae of the right coronary artery reinforce blood supply to this area (10, 11) (see Figure 2). The anterior radiation of the left bundle branch is chiefly supplied by perforating arteries from the left anterior descending coronary artery (13). The posterior radiation of the left bundle branch receives its blood supply from the posterior descending artery of the right coronary artery but also receives some branches of the distal left circumflex coronary artery in the surrounding myocardium (13).

It can be concluded that there is great potential for collateral coronary circulation to supply the bundle branches. Although the main blood supply to the proximal A-V conduction system may be the A-V nodal artery and the first anterior septal perforating branch, the first few septal perforating arteries of the posterior descending branch, the second septal branch of the left anterior descending, the crista branches from the right coronary, as well as Kugel's artery and other atrial branches all provide important secondary collateral channels (12).

The size and distribution of these arteries become critical when considering a narrowing or occlusion of one or more of the coronary arteries in acute myocardial infarction. Even though the definition of blood supply to a given structure must consider the collateral as well as the dominant blood flow, loss of the dominant supply may cause serious impairment to conduction even though myocardial cell death may not ensue.

ACUTE MYOCARDIAL INFARCTION AND COMPLETE HEART BLOCK

There has been much recent interest in the association of complete heart block with acute myocardial infarction, particularly with regard to the indications for pacemaker implantation (16–24). After considering the coronary anatomy as described above, it is clear that the site of the myocardial infarction is important in determining type and severity of block, i.e. anterior or inferior myocardial wall infarction as an indication of insufficient blood supply to the conduction system. In addition to ischemia of the conduction tissue and the degree and the site of the A-V block, the rate of the escape rhythm and the hemodynamic status of the patient must be considered.

Sutton & Davies studied hearts histologically from 24 patients who had inferior wall myocardial infarctions, and in whom the occluded coronary artery had given off the A-V nodal artery (14). The A-V node was entirely normal in 22 cases and showed only focal necrosis in 2. They also studied 5 patients with anterior infarction and found extensive damage to both bundle branches in these hearts. Their total group with acute infarction and heart block numbered 55 patients. If the QRS duration was prolonged, the mortality was 77% (17/22). If the QRS was normal in duration, the mortality was 24% (8/33) (14). Stock & Macken agreed that the mechanism of complete heart block was different in posterior or inferior infarctions as compared to anterior myocardial infarctions (15). In posterior infarctions injury to the A-V node occurred while in anterior infarctions bilateral bundle branch block was seen with necrosis and extensive bundle branch destruction (2, 15) (see Figures 1 and 2).

Table 1 summarizes the type of intraventricular block that occurred with acute myocardial infarction, the incidence of complete heart block, and the mortality. The incidence of bundle branch block or QRS prolongation in acute myocardial infarction varies from 2.6–20% in this series of reports (16–24). There were 408 patients with QRS prolongation and 182 died giving a mortality rate of 45%. The incidence of complete heart block was 26%. Not all of these patients had the onset of their bundle branch block exactly documented. However, this may not be an important factor if the pathology is subjunctional in the bundle branches and is the result of chronic coronary insufficiency (2). What seems important is whether the coronary artery disease is so extensive that the dominant coronary artery circulation as well as the collateral supply is obstructed. If heart block is due to septal necrosis, then this would be in keeping with severe coronary disease and might explain the disparity in incidence of complete heart block (26%) and overall mortality (45%). The majority of patients died with marked left ventricular failure, cardiogenic shock, and pulmonary edema, not heart block alone (18–21, 23, 24). Note that the mortality

Table 1 QRS prolongation with acute myocardial infarction[a]

	RBBB	CHB	MORT	RBBB + LAH	CHB	MORT	RBBB + LPH	CHB	MORT	LBBB	CHB	MORT	ICBBBB	CHB	MORT
Gould–1973 (16)										27	3	5(early) 6(late)			
Gould–1972 (17)	5	0	3	14	10	10	11		7				1		1
Atkins–1973 (18)	3	1	1	30	13	10	1	0	0	18	3	4			
Scheidt–1972 (19)	7	2	2	12	4	4				20	2	7			
Scheinman–1972 (20)	18	3	8	23	3	8	4	1	2	31	3	13			
Lim–1971 (21)	2	1	0	3	2	2	3	3	3	2	2	2	4	4	1
Scanlon–1970 (22)				22	5	8	6	0	2						
Norris–1970 (23)	17	5	12	21	5	11				25	2	12			
Godman–1970 (24)	37	9	11	11	6	5				20	6	13	10	9	9
Totals	89	21	37	136	48	58	25	4	14	143	21	62	15	13	11
Percent		24%	42%		35%	43%		16%	56%		15%	43%		87%	73%

[a] Abbreviations: RBBB = right bundle branch block; CHB = complete heart block; RBBB + LAH = right bundle branch block and left anterior hemiblock; RBBB + LPH = right bundle branch block and left posterior hemiblock; LBBB = left bundle branch block; ICBBBB = incomplete bilateral bundle branch block of other types not included above, e.g. alternating right and left bundle branch block; MORT = mortality. All numbers represent patients, and the entire table contains 408 patients. Of these 408 patients, 182 patients died (45%) and 107 developed complete heart block (26%). For discussion see text.

of patients who developed complete right bundle branch block, complete left bundle branch block, and complete right bundle with left anterior hemiblock is almost the same, 43%. Contrast this with the incidence of complete heart block in these same patients—varying from 15–35%. Patients with complete right bundle branch block and left posterior hemiblock have severe left ventricular dysfunction because of the extensive necrosis needed to interrupt both these bundles and yet these patients have a relatively low incidence of complete heart block, 16%. The same logic should hold for other less common types of incomplete bilateral bundle branch block such as alternating right and left bundle branch block. In their total group of 97 patients Scheinman & Brenman found that 92% of those who died had cardiac failure or shock (20).

Table 2 lists the type of acute myocardial infarction associated with QRS prolongation in those series that reported it. Anterior wall infarctions accounted for two thirds of the cases but one third were associated with other electrocardiographic areas, notably the inferior or posterior wall. The observation that inferior wall infarctions can be associated with a prolonged, conducted QRS further supports the concept of a dual coronary supply to the septum (12). This pattern would suggest that the left anterior descending artery is diseased prior to the development of inferior infarction (see Figures 1 and 2). In other words, chronic coronary insufficiency due to obstructive disease in the left anterior descending artery made the blood supply from the posterior descending artery more critical than it otherwise might have been. When the posterior descending blood flow is reduced or entirely blocked in an acute inferior wall infarction, damage to the bundle branches results because the left anterior descending is already diseased and perhaps dependent on the posterior descending artery for collateral flow.

His bundle recordings have been performed in relatively few patients with acute myocardial infarction (25, 26, 34). His bundle recordings allow the division of the

Table 2 Electrocardiographic location of myocardial infarction[a]

	Anterior	Inferior	Both	Unknown	Total
Gould (17)	26	5			31
Atkins (18)	37	10		5	52
Scheidt (19)	10	9			19
Lim (21)	8	6			14
Scanlon (22)	14	9	2	3	28
Godman (24)	44	13	7	4	68
Totals	139	52	9	12	212
Percentage	66%	25%	4%	5%	100%

[a] Some of the series from Table 1 tabulated the ECG manifestation of the acute myocardial infarction where possible, i.e. almost all of these myocardial infarctions accompanied a sort of right bundle branch block. The important point is that QRS prolongation can occur with acute inferior myocardial infarctions (25%). For further discussion see text.

PR interval into two parts; the A-H interval representing A-V nodal conduction time and the H-V interval which is the time over the distal His bundle and Purkinje system. In the 8 patients studied by Rosen et al, four inferior infarctions and one subendocardial infarction had prolongation of the A-H interval indicating block above the tissue generating the His electrogram (25). Three other patients with anterior myocardial infarctions and complete heart block had prolongation of the H-V interval indicating block below the His bundle. Careful postmortem histological study of the conduction system correlated with antemortem intracardiac electrograms in 7 patients were reported by Hunt et al (26). Of these patients, 5 had significant disease in all coronary epicardial vessels. The 2 others had significant two-vessel disease. Three patients with inferior wall myocardial infarction had block above the His potential as did 1 other patient with an old anterior and recent inferior infarction. The 3 patients with anterior wall infarction all had right bundle branch block and block below the His bundle. Three patients died in heart failure and shock, 3 died suddenly despite electrical pacing, and 1 died in asystole. Localization of the block by intracardiac electrogram records were in good agreement with the histological findings in each case. Lichstein et al performed His bundle recordings in 15 patients with acute myocardial infarction and incomplete bilateral bundle branch block (34). The H-V interval was abnormal in 11 patients and 8 of these died while only 1 of 4 patients with a normal H-V interval died. Hospital mortality in this group of 15 patients was 47% and total mortality was 60%, despite the use of pacemakers. Again the most frequent cause of death in these patients was left ventricular failure.

Simon et al reviewed a series of 134 patients with acute myocardial infarction and some degree of heart block (27). In support of the above data, 7 of 8 patients with acute anterior infarction and complete block died. However, first degree A-V block carried no excess risk in either anterior or inferior infarctions. Four of 6 patients (67%) with second degree A-V block and anterior infarctions died compared to 8 of 22 (36%) with inferior infarctions and similar block (27).

Acute anterior wall myocardial infarctions with complete heart block are less common but more lethal than acute inferior wall infarctions with complete heart block (27). Anterior wall infarction in this setting is associated with more diffuse coronary artery disease and severe left ventricular dysfunction. The placement of a temporary pacemaker is accepted therapy for complete heart block in these patients but mortality may not be decreased significantly because these patients' lives are threatened by the pump failure which continues to progress despite correction of the bradycardia. This is in contrast to heart block with a recent inferior wall infarction where pacemaker insertion is frequently not necessary because the rate of the junctional escape rhythm stays above 50 per minute and these patients do not have such extensive myocardial damage (28, 29). In the series listed in Table 1, there were a few patients with inferior infarction who died in sudden asystole and it is conceivable that they could have benefited from artificial pacing, but the outcome in most of these patients is that they return to normal sinus rhythm. In addition to the effect on rate and rhythm, the presence of an artificial demand pacemaker may facilitate the administration of drugs in these critically ill patients. Finally, once ventricular demand pacing is started, the potential loss of atrial contraction to the cardiac output in these critical patients merits consideration.

Long term follow-up of patients with inferior or posterior wall myocardial infarctions with high grade or complete A-V block, even if the escape rhythm has a widened QRS complex, is no worse than other types of infarctions without left ventricular failure (30, 31). Waugh et al described a high risk subgroup of patients with A-V and intraventricular block with acute myocardial infarction (32). At highest risk for heart block were: (*a*) 12 patients with left bundle branch block; (*b*) 9 patients with right bundle branch block and left anterior hemiblock—both of these groups with additional first degree A-V block; (*c*) 14 patients with right bundle branch block and left posterior hemiblock; and (*d*) 2 patients with alternating right and left bundle branch block. Of the 37 patients, 13 (35%) progressed to complete heart block via a Mobitz Type II A-V block. The risk of complete heart block in hospital increased from 35 to 50% if the blocks were documented consequences of acute myocardial infarction. Twenty six of the 37 patients were discharged without an implanted pacemaker. Of these 26 patients, 10 died suddenly and 2 more experienced syncope with the one-year follow-up period. On the basis of these findings, the authors recommended prophylactic pacemakers for this high risk subgroup of patients (32). Hemodynamics are still important in these survivors of acute myocardial infarction with QRS prolongation. Hamby et al studied 24 such patients and found the left anterior descending coronary artery diseased in all (33). In addition, patients with left bundle branch block and right bundle branch block with left anterior hemiblock had abnormal left ventricular function and disease of two or three coronary vessels (33).

CHRONIC HEART BLOCK

Coronary heart disease does not account for all cases of chronic heart block. In a series of 62 histologically studied hearts, Lenegre found coronary atherosclerosis in only 32 (35). Most chronic heart block is the result of bilateral bundle branch block, not A-V nodal block (2, 3, 5, 6, 8, 35, 36). The etiology of this bundle branch destruction and fibrosis includes coronary artery disease, idiopathic bilateral bundle branch fibrosis, cardiomyopathy, calcific aorta valve disease, myocarditis, connective tissue disorders, and amyloidosis (35, 36). In Lenegre's series of 62 patients, 19 cases had no etiology for the bundle branch fibrosis (35). Davies showed no cause for this fibrosis in 46 of 100 cases (36). Of the 162 cases studied histologically by these authors, 65 had bundle branch fibrosis of unknown cause (40%), while 47 had fibrosis in association with coronary artery disease (29%) (35, 36). Thus, the most common cause of bilateral bundle branch block is unknown. Titus has recently reviewed the correlation between clinical electrocardiographic arrhythmias and pathological changes in the conduction system (37). His bundle recordings in patients with electrocardiographic criteria for bilateral bundle branch block may help to indicate which of these patients will progress to complete heart block (38, 39). A prolonged H-V interval in the presence of bundle branch block suggests disease in the contralateral bundle branch. The majority of patients followed after electrophysiologic studies for bilateral bundle branch block and who develop complete heart block had prolonged H-V intervals recorded in their initial studies (39, 40). Scheinman et al suggest demand pacemakers for patients with neurologic symptoms, bilateral bundle branch block in the ECG, and H-V intervals in excess of 80 msec (41).

FUNCTIONAL ATRIOVENTRICULAR NODAL BLOCK—THE GAP PHENOMENON

Mendez & Moe have recently published a review of the physiology of impulse propagation across the A-V node (42). Damato and his co-workers have recently reviewed the physiology of the A-V node in humans with His bundle recordings (8, 9, 43).

A most interesting type of functional A-V block has received much attention recently—the phenomenon of the gap in A-V conduction (44–49, 51). This was initially described by Moe et al in dog experiments (50). Later it was also found in human hearts by the His bundle recordings (44, 51, 52). The premature or extra atrial stimulus technique was used. Premature atrial stimulations introduced earlier and earlier in the cycle will develop a longer and longer A-V nodal conduction time. If the coupling interval of the premature atrial stimulus to the preceding atrial beat (A_1A_2) is graphed against the coupling cycle of the premature ventricular beat to the preceding ventricular beat (V_1V_2), both cycles tend to shorten (see Figure 3). When the coupling interval of the atrial stimulus is made even shorter, eventually it will not conduct to the ventricles because the A-V node is still refractory from the preceding beat. A blocked atrial premature beat results (see *X*s in Figure 3). If the coupling interval of the premature atrial stimulus is made shorter still, greater delay will be introduced in the A-V node. If the delay introduced by the very premature atrial stimulus is great enough, it may allow more distal conducting tissues time to recover excitability. If the distal conducting tissues recover excitability, the very early atrial stimulation may again conduct to the ventricles. This could result in a much longer ventricular coupling cycle between the premature ventricular beat and the preceding conducted ventricular beat (V_1V_2). It will create the "gap" in A-V conduction (44, 48, 52). This will appear to be supernormal conduction since an earlier premature atrial stimulus conducted to the ventricles while a later one did not (53). The later stimulus did not conduct because all of the distal conducting system was still refractory from the preceding beat. The earlier stimulus did conduct because the delay introduced in the A-V node by the very premature beat allowed sufficient time for the distal system to recover excitability and thus, this is not supernormal conduction (53). The open circles in Figure 3 represent aberrant intraventricular conduction which may be seen on both sides of the gap.

Two conditions are necessary to develop a gap in A-V conduction: (*a*) the absolute or effective refractory period of the distal tissue is longer than the functional refractory period of the proximal tissue; (*b*) the functional refractory period of the proximal tissue must be relatively refractory enough to delay the oncoming premature impulse so that the distal tissue's effective refractory period is exceeded (48). Then the distal tissue can allow conduction to the ventricles. These relationships in refractory periods in these tissues can be upset by increasing or slowing the driving

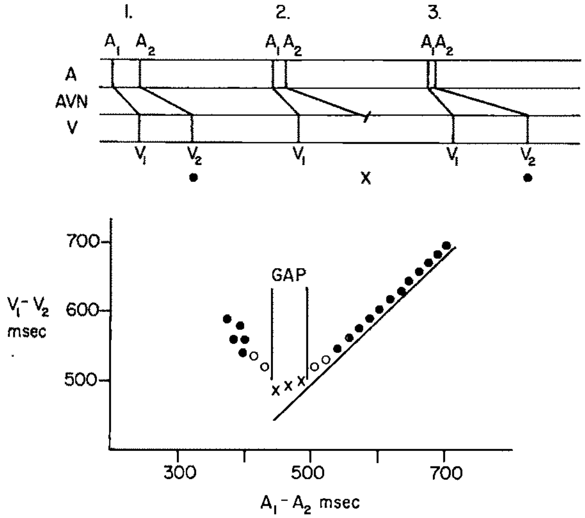


Figure 3 The top part of this drawing is a ladder diagram where A_1 is the last beat of a regular atrial rhythm and A_2 is the premature atrial stimulus. 1. At a certain coupling interval (A_1–A_2), the premature atrial stimulus results in a ventricular beat (V_2) with a V_1–V_2 coupling interval. This premature beat V_2 occurs with an increase in A-V conduction time (A_2V_2). 2. The premature atrial stimulus A_2 comes earlier in the cycle, i.e. at a shorter A_1–A_2 coupling interval and fails to conduct to the ventricles because the His-Purkinje system is still refractory from the preceding beat A_1. This results in a blocked premature atrial beat. 3. At the shortest coupling interval (A_1A_2), A_2 again conducts to the ventricles but with a large increase in A-V conduction time (A_2V_2). This increase in A_2V_2 allows for the recovery of more distal tissues in the His-Purkinje system. Therefore, these tissues allow A_2 to conduct to the ventricles (V_2).

The bottom portion of the drawing is a graph of the atrial (A_1–A_2) and ventricular (V_1–V_2) coupling cycles for the premature beats. The black dots to the right of the gap are in keeping with illustration number 1 above. Gradual shortening of the A_1–A_2 coupling cycle results in a gradual shortening of V_1–V_2. The nonconducted atrial stimuli (*X*) are illustrated by number 2 above. These nonconducted beats cause a gap in A-V conduction. The black dots to the left of the gap are shown by number 3 above. These beats A_2 are much earlier, i.e. shortest A_1–A_2, and conduct to the ventricles with a very long A_2V_2. This long A_2V_2 allows the V_1–V_2 cycle to return to the range where conduction is possible. The open circles represent aberrant intraventricular conduction which can occur on both sides of the gap. See text for a more detailed discussion.

rate or the basic atrial rate (44, 48). The gap will only appear at certain basic cycle lengths. The gap phenomenon then is probably functional and dependent on differences in refractoriness of proximal and distal conduction tissues (46). As further evidence of this, Akhtar et al were able to convert subjects from a Type I to a Type II gap with atropine administration (46).

The gap phenomenon has been categorized as Type I, II, or III depending on the tissues involved proximally and distally. If the tissue that developed the proximal delay was the A-V node and the distal tissue was the bundle of His, this was called a Type I gap (45, 49). Type I gap usually conducts to the ventricles with a normal QRS. Type II gap occurs when the proximal area of conduction delay is in the bundle of His and the distal tissue in the bundle branches (45, 49). It is usually accompanied by aberrant conduction to the ventricles or a widened QRS. The term Type III gap has been used to describe the failure of conduction in the right bundle branch at coupling intervals of 1000 to 610 msec in a patient with left bundle branch block. Conduction resumed at shorter coupling intervals without evidence of additional A-V nodal or His-Purkinje system delay (45). Type II gap shows a delay in the proximal His-Purkinje system when conduction resumes at the shorter coupling intervals (45, 49). The mechanism for Type III is not yet clear and may well represent true supernormal conduction or may possibly reflect the effects of vagal discharges on the conduction tissues (45, 48, 53). However, as Wu et al have pointed out, typing gaps could be an oversimplification (48). Gap phenomena may occur anywhere in the conduction from atria to distal Purkinje cells if the above-listed conditions are fulfilled (48). More recently the gap phenomenon has been demonstrated during retrograde conduction in man (47). In these ventricular to atrial gaps, the initial site of conduction delay can be the His-Purkinje system or the A-V node. These workers demonstrated an antegrade gap in only 1 of 12 patients, while they showed retrograde gaps in 6 of these same 12 patients (47). Again, one tissue slowed conduction sufficiently to allow another more distal tissue to recover excitability.

The gap phenomenon has been used as evidence to support the concept of dual A-V nodal pathways (50). The basic idea of conduction delay induced by a premature stimulus is needed to understand reentry and reentrant supraventricular tachycardia as well as aberrant intraventricular conduction (44, 52). A critical conduction delay in the A-V node or in the His-Purkinje system may result in atrial echo beats, and these echo beats may start the supraventricular tachycardia (54). Reentry can occur in any tissue where there is a difference in conduction velocity and recovery time.

CONCLUSIONS

The pathophysiology of QRS prolongation and actual or impending complete heart block is based on coronary artery anatomy primarily involving the interventricular septum. In complete heart block associated with an acute myocardial infarction, pacemaker implantation may be helpful in selected cases only because of the severity of left ventricular disease in these patients. Perhaps the most important function of pacemakers in this setting will be in association with circulatory support devices. The most common cause of chronic heart block is bilateral bundle branch block caused by idiopathic fibrosis or coronary artery disease. Patients with ECG manifestations of bilateral bundle branch block and a prolonged H-V interval seem to be in particular danger of developing complete heart block. More study and careful follow-up of these patients will clarify the indications for pacemaker implantation here. The A-V gap phenomenon is an example of functional A-V block that is based on differences in refractoriness of different sections of the conduction system. It has added much to our understanding of the physiology of A-V conduction.

Literature Cited

1. Hecht, H. H. et al 1973. Atrioventricular and intraventricular conduction. Revised nomenclature and concepts. *Am. J. Cardiol.* 30:232–44
2. Lev, M. 1973. The pathogenesis of atrioventricular block. In *Cardiac Arrhythmias,* ed. L. S. Dreifus, W. Likoff, 253–57. New York: Grune and Stratton
3. Yater, W. M., Cornell, V. H., Clayton, T. 1936. Auriculo-ventricular block due to bilateral bundle branch lesions. *Arch. Intern. Med.* 57:132–73
4. Rosenbaum, M. B., Lepeschkin, E. 1955. Bilateral bundle branch block. *Am. Heart J.* 50:38–61
5. Lasser, R. P., Haft, J. I., Friedberg, C. K. 1968. Relationship of right bundle branch block and marked left axis deviation (with left parietal or peri-infarction block) to complete heart block and syncope. *Circulation* 37:429–37
6. Lopez, J. F. 1968. Electrocardiographic findings in patients with complete heart block. *Brit. Heart J.* 30:20–28
7. Scanlon, P. J., Pryor, R., Blount, S. E. 1970. Right bundle branch block associated with left superior or inferior intraventricular block. *Circulation* 42:1123–33
8. Damato, A. N. et al 1969. A study of heart block in man using His bundle recordings. *Circulation* 39:297–305
9. Damato, A. N. et al 1969. Study of atrioventricular conduction in man using electrode catheter recordings of His bundle activity. *Circulation* 39:287–96
10. Lev, M. 1968. The conduction system. In *The Pathology of the Heart and Blood Vessels,* ed. S. E. Gould, 180–220. Springfield: Thomas
11. James, T. N. 1965. Anatomy of the coronary arteries in health and disease. *Circulation* 32:1020–33
12. Frink, R. J., James, T. N. 1973. Normal blood supply to the human His bundle and proximal bundle branches. *Circulation* 47:8–18
13. Lev, M. 1964. Anatomic basis for atrioventricular block. *Am. J. Med.* 37:742–48
14. Sutton, R., Davies, M. 1968. The conduction system in acute myocardial infarction complicated by heart block. *Circulation* 38:987–92
15. Stock, R. J., Macken, D. L. 1968. Observations on heart block during continuous electrocardiographic monitoring in myocardial infarction. *Circulation* 38:993–1005
16. Gould, L., Ramana, C. V., Gomprecht, R. F. 1973. Left bundle branch block. Prognosis in acute myocardial infarction. *J. Am. Med. Assoc.* 225:625–27
17. Gould, L., Venkataraman, K., Mohammad, N., Gomprecht, R. F. 1972. Prognosis of right bundle branch block in acute myocardial infarction. *J. Am. Med. Assoc.* 219:502–3
18. Atkins, J. M., Leshin, S. J., Blomqvist, G., Mullins, C. B. 1973. Ventricular conduction blocks and sudden death in acute myocardial infarction. *N. Engl. J. Med.* 288:281–84
19. Scheidt, D., Killip, T. 1972. Bundle branch block complicating acute myocardial infarction. *J. Am. Med. Assoc.* 222:919–24
20. Scheinman, M., Brenman, B. A. 1972. Clinical and anatomic implications of intraventricular conduction blocks in acute myocardial infarction. *Circulation* 46:753–60
21. Lim, H., Toh, C. C. S., Lip Ping Low. 1971. Atrioventricular and associated intraventricular conduction disturbances in acute myocardial infarction. *Brit. Heart J.* 33:947–54
22. Scanlon, P. J., Pryor, R., Blount, S. G. 1970. Right bundle branch block associated with left superior or inferior intraventricular block associated with acute myocardial infarction. *Circulation* 42:1135–42
23. Norris, R. M., Croxson, M. S. 1970. Bundle branch block in acute myocardial infarction. *Am. Heart J.* 79:728–33
24. Godman, M. J., Lassers, B. W., Julian, D. G. 1970. Complete bundle branch block complicating acute myocardial infarction. *N. Engl. J. Med.* 282:237–40
25. Rosen, K. M. et al 1970. Site of heart block in acute myocardial infarction. *Circulation* 42:925–33
26. Hunt, D., Lie, J. T., Vohra, J., Sloman, G. 1973. Histopathology of heart block

complicating acute myocardial infarction. *Circulation* 48:1252–61
27. Simon, A. B., Steinke, W. E., Curry, J. J. 1972. Antrioventricular block in acute myocardial infarction. *Chest* 62:156–61
28. Langendorf, R., Pick, A. 1968. Atrioventricular block, type II (Mobitz). Its nature and clinical significance. *Circulation* 38:819–21
29. Hackel, D. B., Estes, E. H. 1971. Pathologic features of atrioventricular and intraventricular conduction disturbances in acute myocardial infarction. *Circulation* 43:977–79
30. Rotman, M., Wagner, G. S., Waugh, R. A. 1973. Significance of high degree atrioventricular block in acute posterior myocardial infarction. *Circulation* 47:257–62
31. Lie, K. I., Wellens, H. J., Schuilenburg, R. M., Durrer, D. 1974. Mechanism and significance of widened QRS complexes during complete atrioventricular block in acute inferior myocardial infarctions. *Am. J. Cardiol.* 33:833–39
32. Waugh, R. A., Wagner, G. S., Haney, T. L., Rosati, R. A., Morris, J. J. 1973. Immediate and remote prognostic significance of fascicular block during acute myocardial infarction. *Circulation* 47:765–775
33. Hamby, R. J., Tabrah, F., Gupta, M. 1973. Intraventricular conduction disturbances and coronary artery disease. *Am. J. Cardiol.* 32:758–65
34. Lichstein, E. et al 1973. Findings of prognostic value in patients with incomplete bilateral bundle branch block complicating acute myocardial infarction. *Am. J. Cardiol.* 32:913–18
35. Lenegre, J. 1964. Etiology and pathology of bilateral bundle branch block in relation to complete heart block. *Progr. Cardiov. Dis.* 6:409–44
36. Davies, M. J. 1971. *Pathology of Conducting Tissue of the Heart.* New York: Appleton-Century-Crofts
37. Titus, J. L. 1973. Anatomy of the conduction system. *Circulation* 47:170–77
38. Lepeschkin, E. 1964. The electrocardiographic diagnosis of bilateral bundle branch block in relation to heart block. *Progr. Cardiov. Dis.* 6:445–71
39. Gupta, P. K., Lichstein, E., Chadda, K. D. 1973. Intraventricular conduction time (H-V interval) during antegrade conduction in patients with heart block. *Am. J. Cardiol.* 32:27–31
40. Narula, O. S., Samet, P. 1971. Right bundle branch block with normal, left or right axis deviation. Analysis by His bundle recordings. *Am. J. Med.* 51:432–55
41. Scheinman, M., Weiss, A., Kunkel, F. 1973. His bundle recordings in patients with bundle branch block and transient neurologic symptoms. *Circulation* 48:322–30
42. Mendez, C., Moe, G. K. 1972. Atrioventricular transmission. In *Electrical Phenomena in the Heart,* ed. W. C. DeMello, 263–91. New York: Academic
43. Damato, N. A., Gallagher, J. J., Lau, S. H. 1972. Application of His bundle recordings in diagnosing conduction disorders. *Progr. Cardiov. Dis.* 14: 601–20
44. Wit, A. L., Damato, A. N., Weiss, M. B., Steiner, C. 1970. Phenomenon of the gap in atrioventricular conduction in the human heart. *Circ. Res.* 27: 679–89
45. Agha, A. S. et al 1973. Type I, type II and type III gaps in bundle branch conduction. *Circulation* 47:325–30
46. Akhtar, M. et al 1974. Unmasking and conversion of gap phenomenon in the human heart. *Circulation* 49:624–30
47. Akhtar, M., Damato, A. N., Caracta, A. R., Batsford, W. P., Lau, S. H. 1974. The gap phenomenon during retrograde conduction in man. *Circulation* 49: 811–17
48. Wu, D., Denes, P., Dhingra, R., Rosen, K. M. 1974. Nature of the gap phenomenon in man. *Circ. Res.* 34:682–92
49. Gallagher, J. J. et al 1973. Gap in A-V conduction in man: Type I and II. *Am. Heart J.* 85:78–82
50. Moe, G. K., Mendez, C., Han, J. 1965. Aberrant AV impulse propagation in the dog heart: A study of functional bundle branch block. *Circ. Res.* 16: 261–86
51. Durrer, D. 1968. Electrical aspects of human cardiac activity: A clinical-physiological approach to excitation and stimulation. *Cardiovasc. Res.* 2: 1–18
52. Goldreyer, B. N., Bigger, J. T. 1969. Spontaneous and induced reentrant tachycardia. *Ann. Intern. Med.* 70: 87–98
53. Moe, G. K., Childers, R. W., Meridith, J. 1968. An appraisal of "supernormal" A-V conduction. *Circulation* 38:5–28
54. Varghese, P. J. et al 1974. Intraventricular conduction delay as a determinant of atrial echo beats. *Circulation* 49:805–10

Copyright 1975. All rights reserved

URINARY TRACT SYMPTOMS AND INFECTIONS

❖7168

Jay P. Sanford, M.D.[1]
Department of Internal Medicine, University of Texas Southwestern Medical School,
Dallas, Texas 75235

Recent progress in the prevention and control of many infectious diseases has brought to the fore those infections which are caused primarily by the indigenous bacterial flora and those which represent hospital-associated infections. Infections of the urinary tract occupy a major position among these types of infection.

The sensation of discomfort on passing urine is experienced by large numbers of patients at some time or other during their lives. In a survey of 2933 women, aged 20 to 64 years (over 86% of the women in this age group residing in the survey area of South Wales), 21.8% of women said that they had had dysuria during the previous year (1). Age had no significant effect on this proportion. Nearly one half of these women had consulted physicians during this time; however, the proportion of women who sought medical attention was significantly higher in the younger age group. If the incidence of urinary tract infections is based upon patient visits to primary care physicians rather than population surveys, a typical study from Great Britain reveals a peak incidence in women in the 20 year age group with approximately 90 physician visits for dysuria per thousand patients per year (2). This figure is remarkably close to the 10% incidence recorded in South Wales (1). The data from South Wales showing that no significant age difference exists in incidence of symptoms, but that older women less frequently seek medical care, suggest that the apparent peak in incidence of urinary tract infections in women in the 20 year age group and subsequent decline may not reflect a true decrease in incidence, but rather that younger women with symptoms of dysuria seek medical advice more frequently than do older women.

---

[1] The studies of the author reported in this review were supported under Research Grant 2 RO1 AI 10388 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

ACUTE URETHRAL SYNDROME (Dysuria Syndrome or Symptomatic Abacteriuria)

It has long been recognized that many patients seen with complaints of frequency and dysuria have "negative" urine cultures. In the now classical study, Gallagher, Montgomerie & North (3) studied a group of patients seen in the offices of eight physicians in Auckland, New Zealand. A qualified nurse obtained a catheterized specimen of urine which was immediately refrigerated and examined and cultured within one hour of collection. During eight months, 130 patients were studied. The striking finding was "insignificant" bacteriuria ($<$10,000 organisms/ml) in 53 persons (41%). Of greater importance to the observations was the demonstration that 41 of these individuals (32%) had sterile urine cultures. They termed this clinical entity as the acute urethral syndrome. Subsequently, O'Grady and associates have suggested the term "symptomatic abacteriuria" (4). A comparison of the presenting symptoms in individuals with bacteriuria and those with the acute urethral syndrome reveals no significant distinguishing characteristics (Table 1). Note that loin pain, costovertebral tenderness, occurred with comparable frequency in both groups.

Significant pyuria ($>$5 WBC/hpf) was present in 10 of the 77 urines from patients with bacteriuria (13%), as compared with 7 of 53 (13%) of patients without bacteriuria.

The participating physicians were asked to assess the sureness of their diagnoses at the time of initial visit. The primary physicians rated their diagnoses of urinary tract infection as certain or highly probable in 79% of patients with bacteriuric infections, but were equally confident in 60% of individuals with the acute urethral syndrome. Thus, distinction between bacteriuric infections and the acute urethral syndrome cannot be accomplished solely on clinical characteristics.

These overall observations have been confirmed in subsequent studies; Steensberg et al (5) found that 23% of the women and 14% of the men with symptoms of

Table 1 Presenting symptoms[a]

Symptom	Patients With Infection (77)		Patients Without Infection (53)	
	No.	%	No.	%
Fever	28	36	8	15
Dysuria	71	92	38	72
Hematuria	18	23	7	13
Frequency	70	91	50	94
Loin pain	26	34	17	32
Lower abdominal pain	49	64	34	64

[a] Data from reference 3.

urinary tract infection did not have bacteriuria, and Fairley et al (6) found that 23 of 66 (35%) of women presenting to general practitioners with symptoms of acute urinary tract infection had $\leq$ 10,000 bacteria/ml of urine. In the group of 23 patients studied by Fairley, 9 had pyuria (>2000 WBC/ml of urine) (39%) and 5 had gross pyuria (>15,000 WBC/ml of urine). O'Grady et al have shown that patients with frequency and dysuria in the absence of bacteriuria fall into two groups: those who were symptomatic and always abacteriuric and those who were symptomatic but sometimes bacteriuric (4). Urinary leucocyte excretion also fell into two patterns: patients who never had pyuria or had <20 WBC/mm^3 and those who had higher levels of pyuria.

Gallagher et al felt that many cases of the acute urethral syndrome were due to infection confined to the urethra and surrounding glands, especially since bacteriuria subsequently appeared in 13 of 46 of the patients who were initially abacteriuric (3). In 28 of 91 patients with the acute urethral syndrome reported by O'Grady, detection and treatment for various conditions outside the urinary tract, i.e. inflammatory disease of the bowel or vaginitis, cured (or strikingly improved) their symptoms.

The etiology of the acute urethral syndrome especially in those patients who do not have pyuria remains unknown. Studies in progress have included searches for viral agents, chlamydiae (TRIC agents), gonococci, wall defective bacteria, etc. In the patients with pyuria, symptoms, and sterile cultures, it is essential to include tuberculosis, fungal disease of the kidneys, and calculus disease in the diagnostic considerations. It remains to be proven but it is likely that patients with the acute urethral syndrome with pyuria will benefit from antibacterial therapy, while those without pyuria will not.

The importance of recognizing this syndrome(s) is that many of these women do not have bacterial urinary tract infection, will be studied extensively, often treated with costly and/or toxic antimicrobial agents, be subjected to various invasive urological procedures, and will be diagnosed as having "recurrent and chronic pyelonephritis."

BACTERIURIA (Symptomatic and Asymptomatic)

Since the classic work of Longcope & Winkenwerder (7) and Weiss & Parker (8), the occurrence of chronic renal disease as a consequence of pyelonephritis has been well recognized. Weiss & Parker asked, are cases of "silent bacteriurias" benign or can they cause subsequent hypertension or renal failure? Should they be treated as intensely as those associated with the clinical picture of pyelonephritis (9)? Based upon numerous observations such as those of Weiss and Parker as well as those of many others, it was postulated that pyelonephritis might be considered as a disease continuum rather than as a series of apparently unrelated episodes (10).

The concept of a disease continuum reflects a retrospective approach which may be applicable to individual experiences. The problem may be better approached from the epidemiological standpoint, as illustrated in Figure 1. The anatomical changes

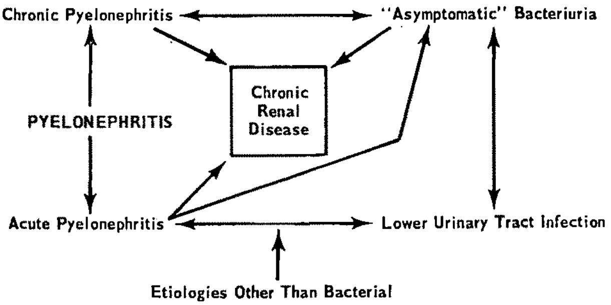

Chronic Pyelonephritis
"Asymptomatic" Bacteriuria
Chronic Renal Disease
PYELONEPHRITIS
Acute Pyelonephritis
Lower Urinary Tract Infection
Etiologies Other Than Bacterial

Figure 1 Kinetic relationship of acute and chronic bacterial pyelonephritis to chronic renal disease.

designated as "pyelonephritis" can be produced by a number of noninfectious etiologies; Freedman presents a list of 14 such causes (11). These will not be considered in this discussion. There is general agreement that both acute and chronic pyelonephritis can lead to the development of chronic renal disease; the unanswered questions relate to the kinetics of these interrelationships, definition of these relationships being paramount in quantifying the problem.

As an initial approach, the assumption that infection limited to the lower urinary tract would not directly cause chronic renal disease, except as a potential source for ascending infection or as an etiology in the genesis of obstructive uropathy, would seem justified. Yet the majority of patients seeking medical care for urinary tract infections do so because of dysuria. How frequently do such patients have bacteriuria and if bacteriuria is present, renal parenchymal involvement?

Localization of the Site of Infection

In their review, Reeves & Brumfitt (12) state: "We consider that a major advance in the understanding of urinary tract infection will occur when a simple and reliable method is available for the localization of infection within the urinary tract of the individual." The methods available for localization of the site of urinary tract infection are summarized in Table 2.

The most reliable method for the localization of the site of involvement has been the recovery of bacteria in ureteral urine obtained from the affected kidney. This technique has been used since the resurgence of interest in the quantitative bacteri-

Table 2 Methods available for localization of infection in the urinary tract

Direct	Indirect
1. Ureteric catheterization	1. Serum antibodies
2. Renal biopsy	2. Urine concentration test
3. Bladder washout technique of Fairley	3. Urinary enzyme excretion
	4. Antibody coated bacteria in urine

ology of urine (13). Stamey, Govan & Palmer (14) have extensively utilized the technique of ureteral catheterization to enable differentiation between pyelonephritis and lower tract infection (Table 3).

Because of the patchy distribution of lesions, renal biopsy is an unreliable tool for estimation of the frequency of pyelonephritis; for example, Freedman performed renal biopsies at the autopsy table on 20 patients with proven pyelonephritis and in only 6 did the biopsies reveal pyelonephritic changes (21).

The indirect method which has been used most widely is that of measurement of antibodies in the patient's serum against the O antigen of the bacteria causing the infection. When the infection involves the renal parenchyma, antibodies are present in significant titers but are absent or present only in low titers when infection is confined to the lower urinary tract. While antibody titers are useful in groups of patients, in individuals there are sufficiently frequent exceptions to make the determination of less value. An example of the correlation is shown in Table 4.

Other investigators have defined renal parenchymal involvement on the basis of impairment of renal function. Winberg reported that 17 of 20 children (77%) with acute nonobstructive urinary tract infections had diminished capacity to produce a concentrated urine (22). This defect repaired in four to six weeks following the

Table 3 Compilation of localization studies utilizing ureteral catheterization

Technique	Ref. No.	Author	(Date)	Type Patient	No.	Sex	Bladder (%)	Renal (%)	Uncertain (%)
Ureteral Cath.	14	Stamey	1965	Referred	95 26	F. M	40 62	60 38	
Ureteral Cath.	15	Whalley	1965	Postpartum bacteriuria	23	F	26	74	
Ureteral Cath.	16	Fairley	1966	Bacteriuria pregnancy	50	F	46	44	10
Ureteral Cath.	17	Fairley	1967	Recurrent infection	70	F	36	64	
Ureteral Cath.	18	Turck	1968	Recurrent infection	51 3	F M	49 0	51 100	
Ureteral Cath.	12	Reeves	1968	Chronic bacteriuria	24	(21 F)	46	54	
Ureteral Cath.	19	Vinnicombe	1971	Referred recurrent	187	(85% F)	36	64	
Ureteral Cath.	20	Whitaker	1973	—	26[a]	(25 F)	73	27	

[a] All patients in this study were children.

Table 4 Comparison of serum antibody titers in patients with renal and bladder bacteriuria[a]

Site of Infection	No. of Patients	Antibody Titer[b]	
		<1:320	≥1:320
Bladder	11	8	3
Renal	13	2	11

[a] Data from reference 12.

[b] Titers determined by direct bacterial agglutination.

response to treatment. In these same children, acidification in response to an acid load and endogenous creatinine clearances were normal in all but three and two children, respectively. Subsequently, Kaitz (23) and Norden & Tuttle (24) reported that 45% and 42% of pregnant women with asymptomatic bacteriuria were unable to concentrate their urine above 700 mosmol/kg of water. Recovery of ability to concentrate their urine returned with successful antibacterial therapy. Again, changes in the ability to produce a concentrated urine are not sufficiently selective to be useful in defining the site of infection in the individual patient.

More recently, Fairley et al (17) have developed an ingenious technique for obtaining ureteric urine without the need to pass ureteric catheters. Their method involves insertion of a triple lumen Foley catheter and collecting a specimen of bladder urine. The bladder is then emptied and 100 ml of sterile saline containing 5 mg of gentamicin (or 2 mg of neomycin) and 125,000 units of topical streptokinase-streptodornase or 2 ampules of Elase® (1 ampule contains 25 units of bovine fibrinolysin and 15,000 units of DNase) is injected into the bladder and allowed to remain for 30 to 45 min. The bladder is again emptied, washed out with 2 liters of sterile water, and the last few milliliters saved for post-washout culture. Urine collections are made every 10 min until five additional specimens are collected. Bacterial counts are done on all specimens. Patients are classified as having lower tract infection if all cultures following the bladder washout are negative. Patients are classified as having upper tract infection if both of the following criteria are met: (*a*) bacterial counts of greater than 10^2 organisms/ml are present in four of specimens 3 to 7; (*b*) there is at least a 10-fold increase between specimen 2 and specimens 3 through 7 (25). Typical results of the bladder washout (BWO) procedure are illustrated:

Site of Infection	*Spec #1 Bladder*	*#2 Immed. post BWO*	*#3 0–10 min*	*#4 11–20 min*	*#5 21–30 min*	*#6 31–40 min*	*#7 41–50 min*
Bladder	>100,000	0	0	0	0	0	0
Renal	>100,000	1000	10,000	80,000	80,000	90,000	90,000
Equivocal	>100,000	400	3600	2800	2000	2400	1800

The BWO procedure has been used quite extensively and correlates well with the results of ureteric catheterization (Table 5). Unfortunately, the BWO procedure is invasive and too cumbersome for routine use.

Thomas, Shelokov & Forland (29) have described a simple technique for the detection of antibody-coated bacteria in urine sediment which appears to correlate with the site of infection. Jones, Smith & Sanford (28) have demonstrated that presence of antibody-coated bacteria correlates very well with renal localization by the bladder washout technique in women; however, in men with prostatic infection antibody-coated bacteria may be detected in the absence of renal involvement (30).

Table 5 Compilation of localization studies utilizing Fairley's bladder washout technique (BWO)

Technique	Ref.	Author	Type Patient	No.	Sex	Bladder (%)	Renal (%)	Equivocal (%)
BWO[a]	17	Fairley 1967	Recurrent UTI	29	F	17	69	14
BWO	26	Bremner 1969	—	150	—	41	59	
BWO	6	Fairley 1971	Acute UTI, gen practice	48	F	46	44	10
BWO	25	Boutros 1972	Prenatal screening	194	F	40	52	8
BWO	27	Gower 1972	—	48	F	38	54	8
BWO	28	Jones 1973	Misc.	26	15F	31	58	11

[a] Correlated with ureteral catheterization. The four equivocal results were in patients with bladder bacteriuria.

Prevalence of Infection by Site

It is standard academic dogma to admonish that knowledge of the natural history of a disease is essential to its rational diagnosis and management. Yet at this time data relevant to the natural history of symptomatic urinary tract infections are both scanty and conflicting. Review of some of the available data will emphasize the necessity for a simple reliable method of localization of infection.

In a study of 66 women who presented to general practitioners with symptoms of acute urinary tract infection, Fairley et al (6) performed cultures and BWO localization studies in those with bacteriuria. The results are summarized in Table 6. Note that one third of patients had the acute urethral syndrome. In this group, one half had costovertebral angle tenderness and 15% actual rigors. Of those with bacteriuria, one half had renal bacteriuria and one half had bladder bacteriuria. Again, differentiation on clinical grounds in any individual patient is not possible.

From these observations, as well as those reviewed previously, the clinical problem of urinary tract symptoms and infection in adults is as follows: Dysuria occurs in approximately 20% of women each year, half of whom seek medical attention. Of these, one third will have the acute urethral syndrome (symptomatic abacteriuria), two thirds will have bacteriuria. Of the two thirds with bacteriuria, one half will have bladder bacteriuria, while one half will have involvement of the renal parenchyma. It is against this background that the available data on the natural history and management of urinary tract infections must be assessed.

URINARY TRACT INFECTION, PROGNOSIS, AND RELATIONSHIP TO CHRONIC RENAL DISEASE

Observations on the prognosis of urinary tract infections (bacteriuria in some instances) in adults are summarized in Table 7. The data from adults preclude generalizations as to the prognosis of urinary tract infections. Most of the studies suggest

Table 6 Acute urinary tract infection in general practice[a]

Manifestations	No Bacteriuria or < 10,000 organisms per ml (23 patients) (%)	Renal Bacteriuria (21 patients) (%)	Bladder Bacteriuria (22 patients) (%)
Symptoms Suggesting Lower Urinary Tract Infection:			
Frequency	95	98	70
Burning	70	68	70
Suprapubic pain	70	68	51
Symptoms and Signs Suggesting Upper Urinary Tract Infection:			
Loin pain	50	48	19
Temperature	35	44	4
Rigors	15	32	15
Nausea and vomiting	25	24	8
Hematuria (macroscopic)	25	20	12

[a] Data from reference 6.

an infrequent progression to chronic renal failure, at least in women. In concurrence with these observations, chronic pyelonephritis was found to be the primary cause of end stage renal failure in 22 (13%) of 173 dialysis candidates, and was almost always associated with an underlying structural defect (42). Unequivocal nonobstructive chronic pyelonephritis was not found. In contrast, the observations of Little, McPherson & de Wardener suggest that acute pyelonephritis is associated with demonstrable destruction of renal mass in three fourths of patients, yet nitrogen retention and focal scarring did not occur (35). The impressions and data of Jackson and associates suggest that one half of patients with active infection for over six years will develop chronic pyelonephritis (32). The view seems inescapable that urinary tract infection causes renal damage in some adults but appears to be infrequent in the absence of significant obstructive uropathy.

On the other hand, there exists much stronger evidence that renal damage occurs more frequently after urinary tract infections in childhood, suggesting that the growing kidney is more susceptible to pyelonephritic damage (Table 8).

Influence of Treatment on Long Term Prognosis

Close review of the prospective clinical trials of Mabeck is of interest (Table 9) (52). All of the patients in this study had two or more cultures with $>10^5$/ml of the same organism on voided urine or a suprapubic aspiration and most had three cultures with $>10^5$/ml. Treatment included a sulfonamide 2.0 g/day or ampicillin 1.0 g/day,

Table 7 Prognosis of urinary tract infection[a] (adults)

Senior Author (Ref.)	No. Patients	Followup (yr)	Course
Freedman (21)	111	1/2–3	47% Rx failure, 0 progression
Pinkerton (31)	50 (BP)	5	21% WBC and bact, 21% with IVP CP
Jackson (32)	—	—	10% of patients with 1+ culture develop CP
	10	6	Active infection throughout 50% CP
Fry (33)	125	4–6	2.4% CP, 7.2% cystitis
Prat (34)	84	1–6	12% had progressive decrease in GFR
Little (35)	16	1/4–2	75% had smaller renal outlines
Whalley (15)	131 (BP)	1/6–1	81% + culture, 27% IVP CP
Bryant (36)	36	>2, av. 36 3.2	36 hypertensive-bacteriurics matched with 36 hypertensive nonbacteriurics. No difference.
Johnson (37)	47	3–6	41 "pyelonephritis", 21 uncleared infection ΔC_{Cr}–12.6, 26 cleared infection ΔC_{Cr}–18.8
Zinner (38)	134	10–14	36 (27%) bacteriuric, + → +: C_{Cr} nl (18): 793 mosmol/kg: IVP-CP 9/32 + → –: C_{Cr} nl (15): 889 mosmol/kg; IVP-CP 3/9
Parker (39)	74	10–20 av. 15.5	11 (15%) had fatal or serious sequelae (unilateral nephrectomy in 3 patients)
Asscher (40)	96	4	BP, urea, renal size — no progression
Freeman (41)	47	25–49 mos.	6 renal failure — 3 infection (2 unrelieved obstruction) 41 no ↓ in renal function

[a] Abbreviations as follows: BP, bacteriuria during pregnancy; CP, chronic pyelonephritis; Rx, treatment; WBC, pyuria; IVP, intravenous urogram; GFR, glomerular filtration rate; C_{Cr}, creatinine clearance; nl, normal limits; bact, bacteriuria.

oxytetracycline 1.0 g/day or chloramphenicol 1.0 g/day for 14 days, with the agent having been selected on the basis of susceptibility tests. Relief of clinical symptoms does not indicate success of treatment. Symptoms disappeared before the bacteriuria was eliminated in 21 of 23 patients on placebo. At both five months and one year there are no significant differences between the groups.

In a long term prospective study, Asscher and associates assessed the results of the treatment of asymptomatic bacteriuria in adults (40, 53). Although symptomatic infections developed more commonly among the bacteriuric than the control women, treatment suitable for use on a large scale did not prevent the emergence of these symptoms. Furthermore, there was no evidence that bacteriuria led to rise in blood pressure, serum urea concentration, or renal scarring.

Since the classic concept of the pathogenesis and course of pyelonephritis is based on the assumption that continuing infection may result in renal failure, long term

Table 8 Prognosis of urinary tract infection (infants and children)

Senior Author (Ref.)	No. of Patients	Followup (yr)	Abnormalities (%)	Comments[d]
Wharton (43)	30	9.6	57	2 in poor health
Woodruff (44)	76	5-23	40	27% + cultures
Macaulay (45)	27	5	41	5 deaths (18%)
De Luca (46)	597	1-10	40.5	Rx acute episodes
Steele (47)	72	11-27	45	11 deaths (15%)
Hodson (48)	175	0[a]	15[b]	Pyelonephritic scarring
Smellie (49)	125	1.4	—	Rx continuous, 11% recurrences, no X-ray abn
	130	1-2	29[c]	84 Rx $\geqslant$ 1 yr — X-rays stable or improved
Lindblad (50)	58F 18M	15-24	19 0	Progressive disease by X-ray
Bergstrom (51)	30	0.5	10	Progressive scarring

[a] Evaluation based upon excretion pyelography of children 0-12 years of age.
[b] Percentage of children with "pyelonephritic scarring."
[c] All 130 children had reflux; 29% had evidence of chronic pyelonephritis.
[d] Abbreviations as follows: Rx, treatment; abn, abnormality; F, girls, M, boys.

therapy has been used by some. The studies of Freeman and associates provide the best data as to the efficacy of continuous chemotherapy (41, 54). At 13 months (54) there was a significant diminution in bacteriuria in the treatment group. However, by 25 months there was no significant difference in the frequency of bacteriuria, although the proportion of patients who developed acute symptomatic exacerbation was less on sulfamethizole. Even the slight gains at 13 months must be balanced against significant drug toxicity in each group.

Harding & Ronald have reported a controlled crossover study of antimicrobial prophylaxis (sulfamethoxazole, methenamine mandelate, trimethoprim-sulfamethoxazole) in 40 women, with each regimen being continued for three months and compared with a three month period with no drug therapy (55). Prophylactic regimens were not started until existing urinary infections had been eradicated. The incidence of infection during the four regimens was as follows: no drug, 3.4 infections per patient year; sulfamethoxazole, 2.5 per patient year; methenamine mandelate, 1.6 per patient year; and trimethoprim-sulfamethoxazole, 0.1 per patient year. In contrast to the study of Freeman et al, drug toxicity was insignificant and did not require cessation of any regimen.

From these studies, it is apparent that if broad concepts are to emerge regarding management of urinary tract infections, localization will be essential. Preliminary studies suggest its value. Turck, Ronald & Petersdorf (18) studied the patterns of recurrence in relation to the site of infection and found that renal bacteriuria was associated with relapse, while bladder bacteriuria was associated with less frequent recurrence of infection, and when it occurred it was usually reinfection, not relapse (Table 10).

Table 9 Course of untreated and treated acute symptomatic uncomplicated urinary tract infection in nonpregnant women[a]

Treatment Regimen (no. Patients)		Follow-Up Interval (1 month)	(5 months)	(12 months)
Placebo (53) →	Untreated (51) →	Sterile (32) Recrudescence (1)	→ Sterile (32) Recrudescence (5) Reinfection (6)	→ Sterile (23) (54%) Recrudescence (10) (23%) Reinfection (10) (23%)
		Persistent bacteriuria (18) →	Recrudescence (5) Treated, sterile (6) Persistent bacteriuria (2)	
	Treated for persistent symptoms (2)			
Sulfonamide or antibiotic for 14 Days (143) →	Sterile during treatment (141) →	Sterile (95) Recrudescence (2) Reinfection (1)	→ Sterile (75) Recrudescence (8) Reinfection (15)	→ Sterile (63) (64%) Recrudescence (9) (9%) Reinfection (26) (27%)
	Persistent infection (2)	Lost to long term follow-up (43)		

[a] Data from reference 52.

Table 10 Correlation between site of infection and pattern of recurrence in chronic bacteriuria[a]

Pattern of Recurrence	Site of Infection	
	Bladder	Renal
	No. of Patients	No. of Patients
Total	25	29
Relapses	11	52
Reinfection	27	13

[a] Data from reference 18.

Evaluation of the optimal duration of therapy as well as type of agent, i.e. one which produces bactericidal urine as opposed to serum levels, remains controversial (14). Some investigators recommend prolonged chemotherapy (56). Numerous other investigators have found no difference in bacteriological cures when a course of two weeks of treatment is compared with a course of six weeks (57–59). Others have even recommended single dose therapy (60). Differences in conclusions may well be the consequence of bias in patient groups based upon differences in localization.

Finally, the importance of localization to therapy is illustrated by the observations of Boutros et al (25) that the instillation of 100 ml of 0.2% neomycin in the BWO procedure eliminated the organisms in 20 of 36 patients (55%) with no recurrences. Observations such as these suggest that the management of bladder bacteriuria may be reduced to single dose therapy while that of renal bacteriuria may require intensive therapy.

From the data developed, it should be apparent that the rational treatment of patients with symptoms suggesting urinary tract infection depends upon: (*a*) differentiation between the acute urethral syndrome and symptomatic bacteriuric syndromes, then (*b*) in patients with bacteriuria, the distinction between bladder and renal bacteriuria. This approach is schematized in Figure 2. When localization

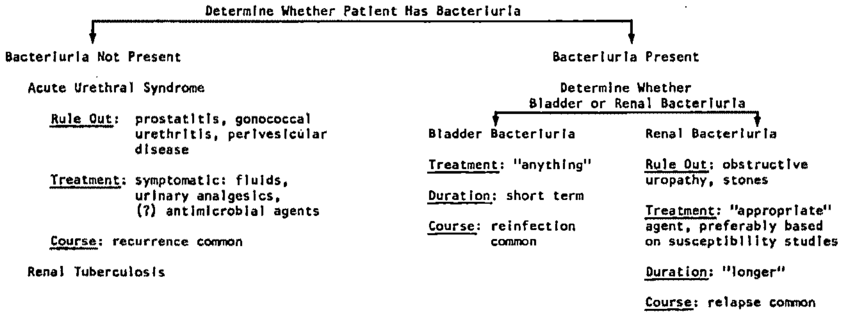

Determine Whether Patient Has Bacteriuria
Bacteriuria Not Present
Acute Urethral Syndrome
Rule Out: prostatitis, gonococcal urethritis, perivesicular disease
Treatment: symptomatic: fluids, urinary analgesics, (?) antimicrobial agents
Course: recurrence common
Renal Tuberculosis
Bacteriuria Present
Determine Whether Bladder or Renal Bacteriuria
Bladder Bacteriuria
Treatment: "anything"
Duration: short term
Course: reinfection common
Renal Bacteriuria
Rule Out: obstructive uropathy, stones
Treatment: "appropriate" agent, preferably based on susceptibility studies
Duration: "longer"
Course: relapse common

Figure 2 Approach to the management of patients with urinary tract symptoms.

of the site of bacteriuria is not feasible, and it should be emphasized that this cannot be done with accuracy on clinical features, bacteriuric patients should be treated as though they had renal bacteriuria.

Correlation Between Site of Infection, Course, and Treatment

In summary, there is general agreement that interrelationships exist between lower urinary tract infection, acute pyelonephritis, chronic pyelonephritis, and chronic renal disease. The kinetics and modifying factors that determine the frequency with which these conditions produce chronic renal disease have been only partially defined. Rational programs directed toward alleviation and control must include longitudinal evaluation based upon adequate localization of the site of involvement.

Literature Cited

1. Waters, W. E. 1969. *Brit. J. Prev. Soc. Med.* 23:263
2. Fry, J., Dillane, J. B., Joiner, C. L., Williams, J. D. 1962. *Lancet* 1:1318
3. Gallagher, D. J. A., Montgomerie, J. Z., North, J. D. K. 1965. *Brit. Med. J.* 1:622
4. O'Grady, F. W., Charlton, C. A. C., Fry, I. K., McSherry, A., Cattell, W. R. 1973. *Urinary Tract Infection,* ed. W. Brumfitt, A. W. Asscher, 81–91. New York: Oxford Univ. Press
5. Steensberg, J., Bartels, E. D., Bay-Nielsen, H., Fanøe, E., Hede, T. 1969. *Brit. Med. J.* 4:390
6. Fairley, K. F. et al 1971. *Lancet* 2:7725
7. Longcope, W. T., Winkenwerder, W. L. 1933. *Bull. Johns Hopkins Hosp.* 53:255
8. Weiss, S., Parker, F. Jr. 1939. *Medicine* 18:221
9. Weiss, S., Parker, F. Jr. 1940. *N. Engl. J. Med.* 223:959
10. Sanford, J. P. 1959. *J. Am. Med. Assoc.* 169:1711
11. Freedman, L. R. 1963. *Diseases of the Kidney,* ed. M. B. Strauss, L. G. Welt, 469–500. Boston: Little, Brown
12. Reeves, D. S., Brumfitt, W. 1968. See Ref. 4, pp. 53–67
13. Sanford, J. P., Favour, C. B., Mao, F. H., Harrison, J. H. 1956. *Am. J. Med.* 20:88
14. Stamey, T. A., Govan, D. E., Palmer, J. M. 1965. *Medicine* 44:1
15. Whalley, P. J., Martin, F. G., Peters, P. C. 1965. *J. Am. Med. Assoc.* 193:879
16. Fairley, K. F., Bond, A. G., Adey, F. D. 1966. *Lancet* 1:939
17. Fairley, K. F., Bond, A. G., Brown, R. B., Habersberger, P. 1967. *Lancet* 2:7513
18. Turck, M., Ronald, A. R., Petersdorf, R. G. 1968. *N. Engl. J. Med.* 278:422
19. Vinnicombe, J., Eykyn, S., Shuttleworth, K. E. D. 1971. *Brit. J. Urol.* 43:39
20. Whitaker, J., Hewstone, A. S., Williams, A. L. 1973. *Invest. Urol.* 10:367
21. Freedman, L. R. 1960. *Biology of Pyelonephritis,* ed. E. L. Quinn, E. H. Kass, 345–53. Boston: Little, Brown
22. Winberg, J. 1959. *Acta Paediat.* 48:577
23. Kaitz, A. 1961. *J. Clin. Invest.* 40:1331
24. Norden, C. W., Tuttle, E. P. Jr. 1965. *Progress in Pyelonephritis,* ed. E. H. Kass, 73–80. Philadelphia: Davis
25. Boutros, P., Mourtada, H., Ronald, A. R. 1972. *Am. J. Obstet. Gynecol.* 112:379
26. Bremner, D. A., Fairley, K. F., O'Keefe, C., Kincaid-Smith, P. 1969. *Med. J. Aust.* 1:1069
27. Gower, P. E., Taylor, P. W., Koutsaimanis, K. G., Roberts, A. P. 1972. *Clin. Sci.* 43:13
28. Jones, S. R., Smith, J. W., Sanford, J. P. 1974. *N. Engl. J. Med.* 290:591
29. Thomas, V., Shelokov, A., Forland, M. 1974. *N. Engl. J. Med.* 290:588
30. Jones, S. R. 1974. *N. Engl. J. Med.* 291:365
31. Pinkerton, J. H. M., Wood, C., Williams, E. R., Calman, R. M. 1961. *Brit. Med. J.* 2:539
32. Jackson, G. G., Arana-Sialer, J. A., Andersen, B. R., Grieble, H. G. 1962. *Arch. Intern. Med.* 110:663
33. Fry, J., Dillane, J. B., Joiner, C. L., Williams, J. D. 1962. *Lancet* 1:1318
34. Prat, V. 1963. *Comptes Rendus du Deuxieme Congress International de Nephrologie,* ed. J. Vostal, G. Richet, 313–14. Publ. House of the Czechoslovak Academy of Sciences, Int. Cong. Series #78
35. Little, P. J., McPherson, D. R., de Wardener, H. E. 1965. *Lancet* 1:1186
36. Bryant, R. E., Windom, R. E., Vine-

yard, J. P., Sanford, J. P. 1964. *Clin. Res.* 12:241
37. Johnson, C. W., Smythe, C. McC. 1969. *S. Med. J.* 62:81
38. Zinner, S. H., Kass, E. H. 1971 *N. Engl. J. Med.* 285:820
39. Parker, J., Kunin, C. 1973. *J. Am. Med. Assoc.* 224:585
40. Asscher, A. W. et al 1973. See Ref. 4, p. 51
41. Freeman, R. B., et al 1975. *Ann. Intern. Med.* In press
42. Schechter, H., Leonard, C. D., Scribner, B. H. 1971. *J. Am. Med. Assoc.* 216:514
43. Wharton, L. R., Gray, L. A., Guild, H. G. 1937. *J. Am. Med. Assoc.* 109:1597
44. Woodruff, J. D., Everett, H. S. 1954. *Am. J. Obstet. Gynecol.* 68:798
45. Macaulay, D., Sutton, R. N. P. 1957. *Lancet* 2:1318
46. De Luca, F. C., Fisher, J. H., Swenson, O. 1963. *N. Engl. J. Med.* 268:75
47. Steele, R. E. Jr., Leadbetter, G. W. Jr., Crawford, J. D. 1963. *N. Engl. J. Med.* 269:883
48. Hodson, C. J., Wilson, S. 1965. *Brit. Med. J.* 2:191
49. Smellie, J. M., Normand, I. C. S. 1968. See Ref. 4, p. 123
50. Lindblad, B. S., Ekengren, K. 1969. *Acta Paediat. Scand.* 58:25
51. Bergstrom, T., Larson, H., Lincoln, K., Winberg, J. 1972. *J. Pediat.* 80:858
52. Mabeck, C. E. 1972. *Postgrad. Med. J.* 48:69
53. Asscher, A. W. et al 1969. *Brit. Med. J.* 1:804
54. Freeman, R. B. et al 1968. *Ann. Intern. Med.* 69:655
55. Harding, G. K. M., Ronald, A. R. 1974. *N. Engl. J. Med.* 291:597
56. Normand, I. C. S., Smellie, J. M. 1965. *Brit. Med. J.* 1:1023
57. Turck, M. et al 1962. *N. Engl. J. Med.* 267:999
58. Kincaid-Smith, P., Fairley, K. F. 1969. *Brit. Med. J.* 1:145
59. Bergstrom, T., Lincoln, K., Redin, B., Winberg, J. 1968. *Acta Paediat. Scand.* 57:186
60. Williams, J. D., Smith, E. K. 1970. *Brit. Med. J.* 4:651

Copyright 1975. All rights reserved

NEUROREGULATORS AND SLEEP MECHANISMS[1]

❖7169

R. Bruce Holman, Ph.D., Glen R. Elliott, B. S., and
Jack D. Barchas, M.D.
Laboratory of Behavioral Neurochemistry, Department of Psychiatry, Stanford University
School of Medicine, Stanford, California 94305

INTRODUCTION

The concept that sleep and waking is controlled by some chemical change dates back to the ancients. Although Piéron (1) proposed the existence of a specific "hypnotoxin" active in brain in the early 1900s, the enormous literature which we have to consider has been built up primarily in the last 15 years. We will not attempt a complete review of this material since many excellent reviews of various aspects of sleep already exist (2–18). Instead, we will restrict our discussion to the most recent neuropharmacological and biochemical data related to hypotheses of the control of sleep by neuroregulators.

Three important advances have been particularly crucial to a scientific approach to sleep. First, Berger's introduction of the electroencephalogram (EEG) meant that sleep was no longer simply an observed state of the organism, but could be directly correlated to electrical changes occurring within the central nervous system (CNS). Moruzzi & Magoun (19) reported that lesions or stimulation within the midbrain reticular system would cause sleep or waking respectively in the cat. This suggested that mechanisms within the midbrain were capable of altering brain function, producing sleep and arousal. However, Aserinsky & Kleitman (20) described two states of sleep based mainly on electrophysiological components. This division of sleep into slow wave (SWS) and rapid eye movement (REM) sleep,[2] makes a simple on-off concept untenable. REM sleep has been further subdivided into what Moruzzi (21)

---

[1] This work was supported by Program Project Grant MH 23,861 from the National Institute of Mental Health and by Award AA 00,498 from the National Institute of Alcoholism and Alcohol Abuse. GRE is the recipient of Fellowship GM 01,922 from the Medical Scientist Training Program, and JDB is the recipient of Research Scientist Development Award MH 24,161 from the National Institute of Mental Health.

[2] Investigators have also referred to SWS as "synchronized sleep," and to REM sleep as "paradoxical sleep" (PS), "desynchronized sleep" (DS), or "active sleep."

has called tonic and phasic events. The tonic events include desynchronized EEG and muscle atonia; while phasic events include a variety of phenomena, most attention has focused on rapid eye movements and ponto-geniculo-occipital spikes (PGO). The PGO spikes are monophasic, sharp waves which are localized to several areas of the brain, including the brainstem, occulomotor nuclei, pons, geniculate nuclei, and occipital cortex. They are associated with eye movements and normally appear just prior to a REM period and throughout REM sleep. In addition, they appear infrequently throughout non-REM sleep, where they tend to occur singly. In REM sleep the PGO spikes occur in bursts. Large single waves are referred to as Type I PGO spikes, while bursts are referred to as Type II PGO spikes. In attempts to determine what specific areas of the brain are involved in regulating these different features of sleep, investigators have shown that lesions, particularly in the brainstem, can selectively suppress waking, SWS, and REM sleep (for a review of this work see reference 22).

Fluorescence histochemistry was the second major development which profoundly influenced the study of sleep. As initially developed by Eränkö, Hillarp, and Falck (see reference 23), this technique has enabled investigators to isolate, visualize, and map specific neurotransmitter cell bodies, their axons, and their terminal endings. Using this technique, it has been possible to develop anatomical maps of the catechol- and indoleamine systems in the CNS. Suddenly stimulations and specific lesions which alter sleep and waking have taken on neurochemical significance.

Third, the work of von Euler, Vogt, Brodie, Axelrod, Carlsson, Udenfriend, their students, and others have led to a new appreciation of the functional mechanisms and regulation of neurochemical processes involving neurotransmitters. Such information has permitted the design of studies which examine the relationship of these processes to sleep. In addition, it has fostered the development of drugs which permit pharmacological dissection of the mechanisms involved. It is this background from which the present-day monoamine and neuroamine theories of sleep have arisen.

We have previously reviewed aspects of an idealized, monamine neuronal model (24, 25). For a nervous impulse to be transmitted from one neuron to the next, a minute quantity of transmitter must be released from the presynaptic nerve terminal ending, diffuse across the synaptic cleft separating the two neurons, and interact with the receptor site to produce a depolarization of the postsynaptic membrane. Synthesis of new transmitter occurs presynaptically to replace the released transmitter. Inactivation of this released transmitter can occur via at least two mechanisms: it may be enzymatically altered, with subsequent removal from the CNS; or, it may undergo reuptake into the presynaptic ending.

Changes in behavior may be related to either an increase or a decrease in neuronal activity. Thus, if a particular transmitter either elicits or inhibits some behavioral function, either at its onset or during its occurrence, one would expect a concomitant change in the amount of transmitter released, catabolized, and synthesized. When these changes are detectable, one can use them to indicate which neurotransmitter systems are either related to or regulating a behavioral state or a change in behavioral state. It is important to emphasize that these changes may be only secondary to the primary regulatory process and that a variety of system interactions are possible.

Two basic strategies are available to the experimenter for defining what and how neurochemical changes in the brain affect sleep and waking: 1. manipulate the neurochemistry of the CNS and observe the effects on sleep patterns, and 2. observe the changes in CNS neurochemistry during various behavioral states. A variety of techniques are available for the first approach. These include lesions, electrical stimulation, and pharmacological treatments. Although they are widely used, the difficulties with each of these methods are their nonphysiological nature and their lack of specificity. These difficulties make the second approach, though technically complex, more attractive, as it deals with a "normal" animal. At present only three methods—cortical superfusion, ventricular perfusion, and push-pull cannula perfusion—have been used for direct measurement of transmitter release in conscious, behaving animals. For the measurement of other biochemical parameters, e.g. synthesis, uptake, storage, and catabolism, animals must be killed at various times during the behavioral state. These are powerful biochemical measures, even though it is difficult to evaluate the influence of factors, other than the behavior of interest, which may alter the biochemical parameters. Despite obvious problems, work relating to the direct neurochemical aspects of sleep is beginning to show interesting results.

There are a wide variety of compounds which are thought to have neuroregulatory properties, either as neurotransmitters or as modulators of neuronal function. Among the compounds for which such a role has been suggested are, in alphabetical order, acetylcholine, γ-aminobutyric acid, cyclic AMP and other cyclic nucleotides, dopamine, epinephrine, glycine and other amino acids, histamine, γ-hydroxybutyrate, 5-hydroxytryptamine, melatonin, a "morphine-like substance," 5-methoxytryptamine, norepinephrine, phenylethylamine and phenylethanolamine, polypeptides, prostaglandins, substance P, taurine, tryptamine, tryptolines, and tyramine. Despite the growing list of compounds with putative neuroregulatory functions, there are only a small number for which there is both strong suggestive evidence of CNS transmitter activity and histochemical localization of their pathways. These three classes of compounds and some references to their histochemical localization are: 5-hydroxytryptamine (5-HT, serotonin) (26–31); catecholamines (CA: norepinephrine, NE; dopamine, DA) (26–28, 30, 32–39); and acetylcholine (ACh) (40–42).

Within the past few years a number of investigators have raised hypotheses regarding the regulation of components of sleep. The most comprehensive and persistent of the various attempts to account for the totality of sleep processes comes from the work of Professor Michael Jouvet (4, 6, 10, 43). To summarize the current monoaminergic theories of sleep: Waking or arousal is regulated by catecholaminergic and cholinergic mechanisms. Specific serotonergic neurons induce SWS; these or other serotonergic neurons trigger the initiation of REM sleep. However, a REM episode is maintained by norepinephrine. Within a REM episode, phasic and tonic events are primarily regulated by serotonergic and cholinergic neurons. Data relating to aspects of sleep regulation will be presented in the material that follows.

5-HYDROXYTRYPTAMINE

Effects of Decreased 5-HT Activity

LESIONS The brainstem raphe nuclei contain the cell bodies of many 5-HT neurons. Destruction of raphe nuclei appears to reduce sleep time and cerebral 5-HT content while increasing waking; however, the effects of lesioning specific raphe nuclei are not consistent, especially with regard to REM sleep.

In the cat and rat, removal of raphe nuclei produces EEG arousal and behavioral waking activity (44, 45). Following 80 to 90% destruction of the cat brainstem raphe, SWS decreases to less than 10 to 15% of the daily recording time; shortly after lesioning, PGO spikes appear in the waking record. Lesser destruction of the raphe complex produces less intense insomnia and more SWS; REM is absent unless the SWS exceeds 15% of continuous recording time. Simon et al (46) also found that destruction of 25 to 40% of the cat raphe increased waking. Lesions of the rostral raphe reduced only SWS, while lesions of more caudal raphe nuclei reduced both SWS and REM. Morgane & Stern (47, 48), using precise lesions, have been able to differentially destroy anterior, middle, and posterior raphe cell groups. They found that only lesions of the anterior nuclei produce a decrease in SWS. Lesions of additional raphe nuclear groups do not produce any further decrease in SWS.

Mancia (49) reported a decrease in sleep both after a midline split restricted to the pons of the cat but including the raphe nuclei, and after cuts lateral to the midline but not directly affecting the pontine raphe. However, no insomnia occurred after destruction of 60 to 70% of the brainstem raphe from midline cuts made in the bulbar and mesencephalic regions.

Jouvet (4) has reported that, in cats, the intensity of insomnia produced by raphe lesions is correlated with the degree of 5-HT and 5-hydroxyindoleacetic acid (5-HIAA) depletion in the forebrain. Rostral raphe lesions appear to decrease both 5-HT and SWS; in contrast, destruction of more caudal nuclei reduces 5-HT content, but produces sleep-onset REM periods which are not preceded by SWS (4, 6, 9). Similarly, Morgane & Stern (47, 48) reported that forebrain 5-HT and SWS are reduced after destruction of the anterior raphe complex. However, they found that lesions of other raphe nuclei neither increase the amount of 5-HT depletion nor further reduce SWS. Raphe lesions in the rat, which produce an increase in waking, also cause a decrease in 5-HT and 5-HIAA (45). These results are further supported by in vitro studies. Tissue from the cortex or brainstem of cats with raphe lesions were examined in vitro and found to have a reduced ability to synthesize and release 5-HT(50).

CHEMICAL LESIONS Recently 5,6-dihydroxytryptamine (5,6-DHT) has been reported to cause relatively specific lesions of 5-HT terminal endings (51, 52). Jouvet and Pujol (53, 54) have reported that an injection of 5,6-DHT into the rostral raphe of the cat produces only a slight, transient sleep change with no change in the 5-HT content. However, injected into the lateral cerebral ventricle, 5,6-DHT produces a significant decrease in 5-HT, 5-HIAA, stage II SWS, and REM sleep; there is a good correlation between the neurochemical changes and the disruption of the sleep states. The preoptic area of the hypothalamus has been implicated in the cortical synchronization of SWS (55, 56). Application of 5,6-DHT to this area is followed by a 60% decrease in SWS and the appearance of behavioral sleep, but with an activated EEG. These findings are consistent with the view that 5-HT preoptic neurons may be involved with SWS cortical synchronization but not with the initiation of behavioral sleep (53).

PHARMACOLOGICAL DEPLETION Reserpine depletes both 5-HT and the catecholamines in the CNS and produces a decrease in SWS and REM sleep in some species (57–59), but an increase in REM in monkeys (60) and man (61–64). As with raphe lesions, reserpine-induced reduction of SWS and REM is generally accompanied by the appearance of PGO activity during waking. This suggests that the drug may not be directly affecting sleep; instead the decrease in sleep may be due to disruption of other neuronal systems, such as those regulating normal PGO activity. Stern & Morgane (65) also reported sleep decrements when reserpine was given intraperitoneally to cats; however, intraventricular injections of reserpine were followed by an increase in REM sleep. Such comparisons are dangerous, but raise the possibility that peripheral effects of the drug may be responsible for some of the behavioral sleep changes it produces. This issue of the peripheral versus central effects of reserpine on behavior may have parallels with other drugs. In addition, reserpine's lack of neurochemical specificity prevents an unambiguous interpretation of its effects. However, studies with reserpine, considered in conjunction with raphe lesion data, lends support to other observations suggesting a connection between 5-HT reduction and the appearance of PGOs outside of REM sleep (46, 66–69).

p-Chloromethamphetamine is more specific than reserpine in depleting 5-HT. In the cat, it causes a marked arousal with PGO spikes in waking and a gradual recovery of sleep (70). *p*-Chloromethamphetamine is thought to deplete 5-HT through inhibition of tryptophan hydroxylase (71); however, 5-hydroxytryptophan (5-HTP), a 5-HT precursor, does not restore sleep in treated cats (70). These findings may relate to Harvey's exciting suggestion that certain chlorinated amphetamines differentially destroy specific raphe nuclear complexes.[3]

p-Chlorophenylalanine (pCPA) is known to block the synthesis of 5-HT by inhibiting tryptophan hydroxylase (72–76), thus depleting CNS 5-HT without affecting the catecholamines. In general, the results of acute injections of pCPA in rats, cats, and primates indicate a close temporal relationship between the cerebral depletion of 5-HT and the suppression of both SWS and REM sleep. Essentially complete depletion of 5-HT results in total insomnia and the presence of PGO spikes in the waking record. Recovery shows a similar correlation between the return of SWS and REM sleep and increases in 5-HT content. Ursin (77) reported that a low dose of pCPA reduces the length and number of deep SWS episodes; reduces the length, but not the number, of light SWS episodes; and has no effect on REM sleep. The pharmacological evidence for the regulation of SWS and absence of waking PGO activity by 5-HT is strengthened by reports that bypassing the enzyme inhibition with low doses of 5-HTP produces a temporary return of normal sleep patterns.

[3] Dr. John A. Harvey, personal communication.

However, several reports provide evidence which is inconsistent with this model. Dement (78) and Jouvet (4) have both reported that PGO spikes are seen in the waking record after pCPA, suggesting that the insomnia may not be "normal" waking. In contrast to the findings of Mouret et al (79), Rechtschaffen et al (80, 81) have reported that, in rats, an 85 to 90% depletion of cerebral 5-HT following pCPA treatment resulted in little or no change in sleep time. Particularly intriguing is the finding that cats kept chronically depleted of 5-HT with repeated injections of pCPA show a return to essentially normal SWS and REM sleep, despite continued low cerebral 5-HT concentrations (82, 83). In man, pCPA does not produce insomnia, although it does decrease REM sleep (84); a slight increase in delta sleep and no change in REM has also been reported (85). It should be noted that the dosages of pCPA given to humans are markedly lower than those used for other animals and that little is known of pCPA's pharmacological effects in humans. Despite the sleep recovery in chronic pCPA-treated cats, PGO spikes continue to appear during wakefulness (82). Henriksen, Dement & Barchas (86) have emphasized the striking parallel between the depletion of 5-HT by pCPA and the appearance of PGO activity in the waking state. This finding again suggests that 5-HT may suppress the PGO activity from SWS and waking (87).

The major problem with the pCPA-derived evidence that 5-HT regulates SWS is the fact that, in the cat, sleep returns despite the continued reduction of cerebral 5-HT content (82, 83). There are several possible explanations for this discrepancy. Pujol et al (50) have found that tryptophan hydroxylase is not completely inhibited after pCPA. Thus, a functional micropool of 5-HT, capable of restoring normal sleep, might be replenished with time. This possibility is given further support by the finding that after pCPA, 5-HT synthesis in the raphe system can be demonstrated with tryptophan loading (76). Alternatively, altered tryptophan metabolism might cause the formation of tryptamine (88) or other metabolites from tryptophan which could serve as 5-HT agonists. Another possibility would be an increase in receptor sensitivity, for which there is indirect evidence. Research in our laboratories has shown that the doses of 5-HTP, the 5-HT precursor, necessary to alter sleep mechanisms are markedly reduced in cats after chronic pCPA treatment.[4] Finally, a change in the biochemical dynamics of another transmitter system may create a new balance which permits reestablishment of sleep. We have found that chronic pCPA administration in cats also markedly depletes norepinephrine, dopamine, and epinephrine.[4] In addition, Stolk et al (89) have shown that NE turnover is reduced in chronically treated rats. It is also of interest that the effects of pCPA can be reversed by tetrabenazine, which depletes catecholamines, or by chlorpromazine (90, 91).

Effects of Increased 5-HT Activity

ELECTRICAL STIMULATION Stimulation of the nucleus medianis of the raphe in the rat has been shown to slightly decrease forebrain 5-HT and significantly increase 5-HIAA (92–94), indicating an increased turnover of 5-HT (95). Aghajanian et al

---

[4] Unpublished observations from the laboratories of W. C. Dement and J. D. Barchas.

(92) did not observe sleep following stimulation; but, more recently, low frequency stimulation of the nucleus dorsalis raphe has been shown to produce a sedative effect (96) with EEG synchronization (93, 97). Synchronization is blocked by lysergic acid diethylamide (LSD), a suspected 5-HT receptor blocker (97). High frequency stimulation at the same site produces arousal and hyperactivity, possibly from current spread to the nearby reticular formation. Gumulka et al (96) found that stimulation of both the dorsalis and medianis nuclei produces an increase in forebrain 5-HIAA, while only stimulation of the dorsalis nucleus is accompanied by sedation.

Responses to electrical stimulation of the raphe are species specific. Stimulation of the ponto-bulbar raphe system in the rabbit elicits arousal (98). In the cat, Jacobs et al (99) found mild arousal or no change in behavior following stimulation of raphe nuclei (see also 100).

ADMINISTRATION OF 5-HT AND RELATED COMPOUNDS Direct application of 5-HT or its metabolites to the brain, either by intravenous administration in animals whose blood-brain barrier is incomplete or via microinjection cannulae directly into the CNS, generally results in sleep or sleep-like behavior (6, 13). Jouvet (4) suggested that a metabolite of 5-HT may have a role in the initiation of REM sleep. Rechtschaffen et al (101) administered 5-HIAA and 5-hydroxytryptophol into the cat brain and found only a very small increase in sleep after 5-HIAA administration. There was no evidence of an effect on sleep or waking either in more recent studies with microapplication of 5-HT metabolites (102) or in studies of the effects of increased metabolites following administration of probenecid to block acid transport (103). Although these data suggest that 5-HT metabolites are not active, careful studies of the full range of 5-HT derivatives must be performed.

The two amino acid precursors to 5-HT, L-tryptophan and 5-HTP, have been given to rats, cats, and man. Hartmann (104) reported that, in rats, a diet containing excess tryptophan does not increase total sleep time, but does increase REM episodes. Acute injection of tryptophan produces only a slight increase in total sleep time and a decrease in REM (105). Radulovacki & Karmos (100) have also reported that L-tryptophan, in doses that increase 5-HIAA content in the CSF, produces no changes in SWS. Large doses of 5-HTP produce cortical synchronization and a sleep-like behavior in cats (43), rabbits (106), and monkeys (107). Differences from normal EEG patterns of sleep make these findings difficult to interpret. In some cases, doses of 5-HTP are reported to elicit excitation or sedation, depending on the time after injection (108, 109). These findings may result either from uptake of 5-HTP into nonserotonergic neurons where it may be decarboxylated to act as a false transmitter or, theoretically, from receptor blockade. In man, 5-HTP produced a decrease in the latency of REM sleep in 5 of 16 patients (110). Other investigators have described an increase in REM sleep with a trend toward decreased SWS (111, 112). Studies along these lines are of particular interest, because 5-HT precursors are of potential use in the treatment of several human disorders.

MONOAMINE OXIDASE INHIBITORS The effect of MAO inhibitors to decrease REM sleep has been a most consistent finding. These compounds act to block one route of 5-HT catabolism. They have been found to produce an increase in cortical synchronization and SWS and a decrease in REM in the cat (113). In the rat, REM sleep is also decreased, but is accompanied by arousal and decreased SWS (114). The effect of MAO inhibition is very complex, for it produces a shift in the pattern of metabolism of many monoamines, including 5-HT. In addition, MAO inhibitors have a multiplicity of effects on the serotoninergic system, including reduction of the firing rate of 5-HT raphe cells (115). Thus it is difficult to definitively relate their effects to changes in 5-HT activity.

Effects of Sleep and Waking on 5-HT Activity

NEUROPHYSIOLOGICAL PARAMETERS Single unit recordings of identified dorsal raphe cells demonstrate reduced firing rates during SWS and again during REM sleep (116). 5-HT may suppress behaviors which are incompatable with sleep; that is, 5-HT may allow sleep to occur by constraining or preventing waking (78). By this view, once sleep has been initiated, it can be maintained without need for 5-HT neuronal activity. However, Bloom et al (117) reported that some 5-HT cells in the raphe magnus show an increase while others show a decrease in activity when unanesthetized cats go from waking to SWS to REM sleep. Some cells in this nucleus showed rapid bursts of activity in REM sleep, possibly associated with PGOs or with bursts of eye movement.

NEUROCHEMICAL PARAMETERS Studies are just beginning to examine biochemical changes and their relationship to sleep states and 5-HT neuronal activity.

Sinha et al (118) killed cats after 10 min of non-REM sleep or 5 min of REM sleep and compared the amine concentrations in various regions of the brain to waking values. The authors reported a significant decrease of 5-HT content in the cerebral cortex with non-REM sleep. This result is difficult to interpret, since there is no known necessary link between tissue content of a neuroregulator and the functional state of a particular neuronal system. However, this marked and unexpected change suggests either a decreased synthesis or an increased utilization of 5-HT which occurs over a matter of minutes. Further studies are required to examine these metabolic possibilities and the kinetics of the process to determine when the changes occur in relation to the sequential pattern of sleep states.

Consistent with a relationship between 5-HT activity and sleep, Mitler et al (119) reported that mice from the C57 strain show more SWS and, at peak sleep periods, more REM sleep than mice from the BALB strain. They also have higher brain stem and cortical tryptophan concentrations, higher tryptophan hydroxylase activity, and higher cortical 5-HT content.

Wyatt and his colleagues (120) reported that, in humans, 5-HIAA concentrations in CSF from ventricular shunts increase during SWS. This apparent increased 5-HT neuronal activity is present only in SWS, not REM sleep. However, in cats no correlation was seen between CSF 5-HIAA content and sleep (121); this lack of correlation may result from the fact that the investigators perfused a large portion of the cerebral ventricular system, thereby potentially masking changes from specific tissue loci.

Investigators have also attempted to determine the neurochemical changes associated with deprivation of REM sleep. The "flower pot technique" permits relatively selective deprivation of REM sleep, and several authors have described increases in 5-HT turnover and synthesis using this paradigm (122, 123). Sinha et al (124) have reported an increase in tryptophan hydroxylase activity following REM deprivation, a finding which is consistent with a change in the dynamics of the 5-HT system. REM deprivation does not appear to increase 5-HT metabolism by a pituitary stress mechanism, since deprivation produces a more rapid and larger increase in 5-HT turnover in hypophysectomized rats (125).

CATECHOLAMINES: NOREPINEPHRINE AND DOPAMINE

Effects of Decreased CA Activity

LESIONS Many investigators have made extensive use of brainstem lesions in attempts to identify the origin of REM sleep control (43). Of interest to this review is more recent work which relates such lesions to the anatomical distribution of the catecholamines. Jouvet (6) has thoroughly summarized the relation between CA neuron destruction and changes in sleep and waking (see also reference 126). Destruction of DA cell bodies in the cat reduces the forebrain DA concentration. The magnitude of these reductions correlates well with the decrease in behavioral activity which also occurs; however, the cortical EEG shows normal waking and sleep patterns. Work in Jouvet's laboratory indicates that NE cell bodies are separable into two groups—those associated with changes in waking behavior and those related to REM sleep. Lesions in the reticular formation reduce cortical EEG activation associated with waking but do not affect behavioral arousal. These lesions produce a decrease in forebrain NE which correlates with a decrease in desynchronization (126).

A marked increase in SWS and REM occurs with lesions of the dorsal NE bundle at the junction of the pons and mesencephalon. These lesions reduce cerebral NE content; in addition, brain slices from these animals are reported to have increased 5-HT turnover (127). The dorsal NE bundle arises from the rostral third of the locus coeruleus, and lesions of this region produce a temporary increase in sleep with no effect on PGOs. More caudal lesions in the locus coeruleus and nucleus subcoeruleus affect tonic and phasic events of REM sleep. Destruction of the most caudal area of the locus coeruleus reduces the NE content of the ventral pons and of the cervical spinal cord; it seems to remove the muscular inhibition generally associated with REM sleep (43). In rostral areas of the midbrain, the degree of REM sleep suppression appears to depend on the extent of the lesion; PGO activity continues to decrease, but is not totally inhibited until bilateral lesions of all the nuclei containing NE are made (128).

Both the difficulties in preparing and maintaining lesioned animals and the complexities of sleep recording procedures make these studies very time consuming. Not surprisingly, there are few attempts to confirm previous studies; however, at least two investigators have attempted to confirm the above findings. Henley & Morrison (129) have confirmed that overt behavior is released into REM episodes following the destruction of the caudal locus coeruleus. However, Jones et al (130) have not been able to replicate the effects on sleep of total locus coeruleus lesions. These authors have found that lesions of the NE cell bodies of the pons of the cat, resulting in 85% depletion of cortical NE, produce no effect on the percentage of time spent in tonic cortical activation and no change in REM sleep, except for a lack of the usual muscle atonia. These animals further failed to show any qualitative or quantitative differences in the duration of arousal following treatment with amphetamine. Jones et al (130) did see a decrease in PGO activity; however, this decrease was only to 50% of normal and occurred gradually over the first week after the lesion, suggesting a secondary cellular degeneration effect rather than a direct NE interaction.

CHEMICAL LESIONS 6-Hydroxydopamine (6-OHDA) chemically lesions CNS catecholamine terminals. Hartmann et al (131) reported that, following the intraventricular injection of 6-OHDA in rats, REM sleep is increased and waking is reduced with decreased cerebral catecholamine content. Direct injection of 6-OHDA into the dorsal NE pathway in the mesencephalon has been shown to reduce cortical NE (39). It also reduces the EEG waking patterns in rats but does not affect behavioral waking; REM sleep is unaffected or slightly decreased; and there is an increase or no change in SWS (132, 133). In cats, intraventricular injection of 6-OHDA has been shown to reduce not only NE and DA, but also 5-HT and 5-HIAA (134). There is an initial excitation and hypothermic condition, followed by sedation and appearance of continuous PGO discharges (135–137). This excitation is inhibited if cerebral NE is depleted by pretreatment with the synthesis inhibitor α-methyl-*p*-tyrosine (AMPT) (135). Prior administration of chlorimipramine protects 5 -HT terminals by stopping uptake of 6-OHDA; under these circumstances, non-REM PGO discharge and hypothermia do not occur (135). With this selective damage of CA terminals, there is a significant decrease in cortical synchronization; REM is normal at 6 days and then decreases; and PGO activity in REM is unchanged. Based on results following selective pharmacological depletion of the amines, Haefely et al (68) further suggested that NE, like 5-HT, produces a tonic inhibition of PGOs in the cat.

PHARMACOLOGICAL DEPLETION AMPT acts to inhibit tyrosine hydroxylase, thus producing a depletion of both DA and NE. Although reports differ, AMPT generally seems to reduce waking in cats (138), and both waking and REM sleep in rabbits (139) and monkeys (140). In the rat, investigators have reported no change (141) or decreases (142) in REM sleep after intraperitoneal injections of AMPT. However, Hartmann et al (143) found an increase of REM sleep when AMPT was administered orally. Inhibition of CA synthesis in rats also fails to affect the normal rebound of REM sleep following REM deprivation (144). Both King & Jewett (145) and Henriksen et al (146) found that cats have increased REM sleep with cerebral NE depletion after AMPT administration (see also 147). Wyatt et al (148) and Vaughan (149) have found that AMPT increases human REM sleep, with no effects on non-REM sleep; REM sleep sometimes returns to baseline within 5 days. α-Methylphenylalanine, which reduces catecholamines in the brain via the same mechanism as AMPT, also increases REM sleep in human patients.

Inhibition of dopamine-β-hydroxylase (DBH) with disulfiram depletes NE and increases cerebral DA. Disulfiram administered to cats results in a decrease both in REM sleep and in the PGO rate during REM; there is also a decrease in waking (150). FLA-63, which also inhibits dopamine-β-hydroxylase, produces a decrease in waking and REM sleep. There is a corresponding increase in SWS which parallels NE depletion; however, by the time NE is maximally depleted, REM sleep has returned to baseline values (151).

α-Methyl-DOPA releases NE from terminal endings and acts as a false transmitter. In the cat, it reduces NE content in the brain and suppresses REM sleep (150). In humans, α-methyl-DOPA only affects REM sleep during the early part of the night, with no change in total REM sleep. This finding is complicated by the possibility that the drug may also decrease 5-HT concentrations in humans (152).

Catecholamine receptor blocking agents reduce neuronal transmission and have been reported to decrease REM sleep in animals (153). Reports in humans indicate that receptor blocking agents of both the alpha and beta type decrease non-REM sleep (154) with no effect on REM sleep (155).

Delorme et al (156) reported that reserpine, given to cats, releases both 5-HT and CA, producing a suppression of SWS and REM and releasing PGOs into waking. Reserpinized animals, treated with dihydroxyphenylalanine (L-DOPA), regain REM sleep and its tonic signs (157, 158); but PGO activity continues (156). In man, reserpine increases REM sleep (61, 62, 64). Amphetamine, which initially releases CA and ultimately depletes them, also elicits marked excitation. This excitation does not occur if CA are first depleted with AMPT (5, 159). The DA receptor blockers, pimozide and spiroperidol, have no effect on amphetamine excitation, implying an NE effect (160). Large doses of amphetamines administered to humans produce excitation and can lead to a psychotic condition; low doses produce a decrease in REM sleep (161, 162). The reuptake inhibitors, imipramine and desmethylimipramine, also decrease EEG waking and REM sleep in cats (6, 163). Similar results have been reported in humans (164, 165). However, Holman et al (166) have shown that clorgyline, a drug which is thought to directly stimulate NE receptors, produces sedation and a synchronized EEG when given intraventricularly to rats.

Effects of Increased CA Activity

ELECTRICAL STIMULATION The role of the ventral NE system in self-stimulation behavior has been of great interest. Stimulation of this area produces behavioral activation and an increased turnover of NE (167); but, as we have already noted, lesions here do not affect waking or sleep. Stimulation of the dorsal NE system produces marked behavioral arousal (19).

ADMINISTRATION OF CA AND RELATED COMPOUNDS Microinjection of NE into the 4th ventricle of the cat produces arousal (168). NE, applied directly to the preoptic nucleus of the hypothalamus or the reticular activating system of the rat during deep sleep, reduces the latency and duration of REM (142). Torda (142) has also shown that, during a REM episode, NE reduces the latency of arousal. She believes that these results are compatible with an NE involvement in arousal.

Drugs which increase the CNS content of CA generally increase arousal. As a precursor of DA and NE, L-DOPA alone or in conjunction with a peripheral decarboxylase inhibitor causes persistent arousal in the cat (169), although this has not always been confirmed (170). Reis et al (171) reported that cats treated with L-DOPA show a correlation between the degree of excitation and brainstem NE content. The majority of studies with humans have shown an increase in REM sleep, but some investigators have reported no change or decreases in REM sleep (155); no effect has been seen on non-REM sleep, although high doses of L-DOPA produce insomnia (172, 173). Unfortunately, the uptake and conversion of L-DOPA is not specific to catecholamine neurons (174), nor is it possible to differentiate between DA and NE effects.

Dihydroxyphenylserine (DOPS), although not a naturally occurring amino acid, can be directly decarboxylated only to NE. In rats, DOPS increases both SWS and REM (175); but it has no effect on sleep in man (176). However, there is disagreement as to whether DOPS crosses the blood-brain barrier (155).

MONOAMINE OXIDASE INHIBITORS The increase in cerebral NE concentration which follows MAO inhibition is accompanied by a marked decrease in sleep, especially REM sleep in the rat (114) and cat (4), and an increase in cortical desynchronization in the rabbit (108). Similarly, in humans, high doses of phenelzine, mebamazine, and pargyline result in up to a 100% decrease in REM sleep (112, 148). Generally non-REM sleep increases, so that there is no change in total sleep time. As we have already noted, MAO inhibitors increase the content of both 5-HT and CA, making it difficult to interpret these results.

Effects of Sleep and Waking on CA Activity

NEUROPHYSIOLOGICAL PARAMETERS Two groups have examined the activity of single units in the locus coeruleus nucleus. Chu & Bloom (177) found NE units in which activity was reduced with "drowsiness" and SWS, but showed bursts of firing during REM. Bubenick & Monnier (178) found that the nuclear mass diameter of locus coeruleus cells in the rabbit decreases in size during thalamic-stimulation-induced sleep, as compared with waking—changes which are believed to be indicative of alterations in neuronal activity.

NEUROCHEMICAL PARAMETERS There have been few reports of biochemical measures related to CA during sleep or after the manipulation of sleep. Rats deprived of REM sleep show an increase in the synthesis and "disappearance" of NE during recovery from deprivation (179). However, non-REM sleep is reduced as well, making it difficult to decide if the neurochemical changes result from REM rebound or recovery of non-REM sleep. Sinha et al (124) reported an increase in the cortical content of NE following 5 min of REM sleep. No difference was observed between waking and non-REM sleep. These results may indicate either an increase in NE synthesis with no concomitant increase in release, or a decrease in NE release.

ACETYLCHOLINE

Effects of Decreased ACh Activity

Procedures which elicit cortical synchronization, including lesions (180, 181), sleep (182, 183), and barbiturate anesthesia (182) also decrease the cortical efflux of ACh. Interestingly, anticholinergic drugs, in cats, produce waking behavior accompanied by electrocortical patterns which resemble SWS (184). Cholinergic blocking agents, such as atropine, produce a similar effect in dogs (185); particularly striking is the similarity of the sleep spindles and slow waves of the drug-induced EEG to those seen in control records. Given to normal and pontine-transected cats, atropine also selectively suppresses REM sleep (186). In humans, scopolamine is reported to decrease REM sleep and increase non-REM sleep (187).

As previously mentioned, pCPA administration is accompanied by the occurrence of PGO spikes in waking and SWS. Jacobs et al (188) reported that catecholamine receptor blocking agents have no effect on these pCPA-elicited PGOs; in contrast, very small doses of atropine are effective. This blockade of PGO spikes occurs even in deeply barbiturized cats, suggesting that the atropine effect is not due merely to increased excitation. PGO activity markedly increases following REM deprivation (189). Henriksen et al (190) found that, following REM deprivation, atropine prevents the appearance of Type II PGOs but does not abolish Type I PGOs.

Effects of Increased ACh Activity

Procedures which elicit cortical arousal, including lesions (191, 192), electrical stimulation (193, 194), and amphetamine administration (195) produce an increase in the cortical release of ACh. Application of ACh via the carotid artery increases cortical activation (196). Other cholinergic agonists also increase both behavioral and cortical arousal (197–200). With nicotine, this arousal is blocked by lesions of the mesencephalic reticular formation (200), suggesting a brainstem site of action; however, other investigators have suggested that cholinergic agonists produce direct cortical activation (201).

Direct chemical stimulation of the ascending cholinergic system in the brain produces a significant increase in waking time. Depending upon the site, microapplication of ACh in the brain can produce sleep as well as arousal (202–207). Recently Sharpe et al (208) found that an injection of carbachol into the dorsal raphe of the monkey produces total insomnia. However, the same compound, administered in the locus coeruleus, increases SWS and total sleep time. Cordeau et al (202) and George et al (209) reported that cholinergic agonists, injected into the region of the pontine reticular formation of cats, produce a REM-like state. Using a similar paradigm, Mitler & Dement (210) found motor inhibition with PGO spiking. Atropine blocks these effects (210). These results further suggest that ACh neurons in the raphe and locus coeruleus may have a role in the inhibition or cessation of normal 5-HT and NE activity.

Cholinesterase inhibitors prevent the catabolism of ACh. In reserpinized cats, eserine induces EEG and muscular changes very similar to those seen in REM sleep, suggesting a cholinergic involvement in REM (211). 2-Octyl-γ-bromo-acetoacetate has also been shown to increase REM sleep in encephale isolé cats; this effect may be related to its anticholinesterase activity (212). Similar reports in man suggest that increases in REM sleep occur with exposure to anticholinesterases (155). Pompeiano and his collaborators (213) have reported that eserine elicits bursts of eye movements and Type II, but never Type I, PGOs. The effects of eserine are blocked by lesions of the medial and descending vestibular nuclei. These last findings agree well with the effects of atropine on PGO activity.

Effects of Sleep and Waking on ACh Activity

Sleep produces a decreased ACh content in samples collected from superfusion of the cerebral cortex (182, 183). The possible involvement of ACh in REM sleep is supported by reports that REM sleep deprivation produces a decrease of bound ACh in the telencephalon (214, 215). However, the excitation which occurs during REM deprivation has not been ruled out as a compounding factor. Jasper & Tessier (216) have reported that the cortical release of ACh is higher during REM sleep than SWS. These results may reflect differences in cortical electrical activity between these states. Push-pull cannula studies of the cat striatum provide similar neurochemical results (217). Ventricular perfusates from conscious dogs indicate increases in ACh release during REM sleep and just prior to waking (218).

OTHER POSSIBLE NEUROREGULATORS

Although these three amines have been the focus of most sleep research, there are numerous other substances in the brain whose relation to sleep has yet to be thoroughly examined.

γ-HYDROXYBUTYRATE γ-Hydroxybutyrate (GHB) is a metabolite of γ-aminobutyric acid (GABA) (219, 220). It has a dose-dependent sedative or anaesthetic effect, and low doses of GHB produce an onset of REM sleep in cats (221, 222). However, this was not confirmed by Winters & Spooner (223), who found no change in sleep with low doses of GHB and an epileptic EEG record with high doses. Similar results have been reported in humans: coma followed large doses of GHB, but there was never a consistent sleep change (224). In rats, GHB results in an increase in DA (225), with no effect on GABA, NE, or 5-HT content in the brain; however, in reserpinized rats, GHB still elicits sleep (226).

HYPNOTOXINS Finally, we return to Piéron's "hypnotoxin" and its more modern counterparts. The experiment of Legendre & Piéron (227) and its replication by Schnedorf & Ivy (228) involved the removal of cerebrospinal fluid (CSF) from sleep-deprived dogs and its injection into normal recipients. Although both these studies reported sleep induction, the stress of the experimental manipulations leaves the conclusions in doubt.

Monnier and his colleagues (229) first reported that stimulation of the mediocentral intralaminary nucleus of the thalamus in rabbits produces an increase in the delta rhythm of sleep with the appearance of "Sleep Factor Delta" in the circulation. Cross perfusions of blood between a stimulated donor and a non-stimulated recipient produce a similar delta-activity increase in the recipient. In a related paradigm, stimulation of the midbrain reticular formation produces a decrease in delta activity in both animals (230). These original reports have been followed by a series of papers in which Professor Monnier and his colleagues have further attempted to isolate and identify the postulated hypnogenic and arousal factors (231–237). A dialysate of cerebral venous blood from rabbits in stimulation-induced sleep, when injected into a normal recipient, produces a state which, by EEG and behavioral criteria, does not differ from physiological sleep (230). The most recent reports on Sleep Factor Delta have suggested that it is a low molecular weight polypeptide; the molecular weight appears to be between 350 and 1500 with the most active fraction being around 700. Further examination of a "sleep dialysate" in rabbits indicates that its sleep-producing effects are not due to changes in arterial blood pressure, CSF pressure, or respiration rate. As it does in normal sleep, bradycardia accompanies infusion and follows the course of delta activity, suggesting correlation with, rather than causation of, sleep.

If one grants the existence of a humoral "sleep-inducing substance," one might expect to see increased concentrations of this substance when the demand for sleep is increased by deprivation. Ringle & Herndon (238, 239) have found that dialysates of plasma or CSF from sleep-deprived rabbits, when injected into normal rabbits, rats, or mice, did not produce sleep. It is not readily apparent why this sleep-deprivation paradigm failed to work.

This failure is especially unexpected in light of work by Pappenheimer and his associates (240, 241). Their original report described the behavioral effects of CSF samples from sleep-deprived goats, when injected into the cerebral ventricles of cats and rats. The results, in both species, are a reduction in behavioral activity and initiation of a sleep-like state. This sleep-like state has been shown to resemble natural sleep by EEG analysis. This "Factor S" is reported to be of low molecular weight (less than 500), and to appear in CSF only during sleep deprivation. The effects of this substance cannot be duplicated by infusions of 5-HT, r-hydroxybutyrate, butyrolactone, GABA, or glutamic acid, even when these compounds are injected in concentrations ten times greater than their normal CSF concentrations. In rats, a long-lasting hyperactivity follows infusion of a high molecular weight dialysate fraction from the CSF of either sleep-deprived or control goats.

Haulica et al (242) found that the cerebral content of adenosine is increased in rats deprived of REM sleep for 48 hr. Intraventricular injections of adenosine in dogs produce EEG and behavioral SWS, raising the possibility that adenosine may be identical to Factor S or Sleep Factor Delta.

Finally, Drucker-Colin and his associates (243, 244) have studied the effects of push-pull perfusates from specific tissue sites in the midbrain reticular formation of sleeping cats. When injected into the same brain loci of a recipient cat, these perfusates produce an increase in the duration and a decrease in the latency of SWS. As yet, the responsible neuroregulator has not been identified.

These studies of hypnotoxins pose exciting possibilities. However, many problems still exist with this work, including the nonphysiological production of sleep. Further, the use of ruminants, whose sleep patterns differ markedly from nonruminants, raises the possibility that these hypnotoxins are species specific. Assuming that the existence of these compounds can be verified and their structures determined, an investigation of their mode of action and relation to other sleep mechanisms is of utmost importance.

SUMMARY

Information which has emerged thus far relates to the overall transmitter mechanisms of sleep. The data, while conflicting, point to the involvement of many neuroregulators at numerous integrative levels of the process. However, the long term questions still remain: what triggers and maintains sleep, what stops sleep, what occurs to the body and brain during sleep—in essence, why sleep? These questions are now problems for behavioral neurochemists, whereas in a previous era, they were problems for philosophers. Unfortunately, our answers to date, while in another idiom, have hardly been more complete or satisfying.

To answer these questions, it will be necessary to understand, in detail, the manner in which neurobiochemical processes relate to the functional physiology of sleep. Although existing studies have given invaluable insight into the neurochemical anatomy of sleep, we have only recently acquired the technical and biochemical expertise necessary to investigate sleep as it occurs normally. Future research must focus on the dynamic changes associated with the regulatory mechanisms of neurotransmitters. Many questions can be asked. With sleep transitions, what changes occur in transmitter content, synthesis, or release? Are there changes in metabolic pathways, reflecting a shift from intra- to interneuronal metabolism? What changes occur in pre- and postsynaptic neurotransmitter receptors to affect sensitivity? What constraints do genetic (245) and environmental (246) factors impose upon these mechanisms? Knowledge of such parameters will allow us to construct more complete models of the neuroregulatory basis of sleep and waking.

However, as we acquire this knowledge, we must avoid the temptation of assuming causation when the evidence merely shows correlation. Neuroregulators are involved in the control of a number of different behaviors; and, at present, we have few, if any, methods for establishing causative links between a specific neuroregulator and a specific behavioral state. Yet, even without an understanding of what "causes" sleep, we may be able to develop pharmacological agents which permit discrete alteration of sleep mechanisms in a more physiological and specific manner.

This potential for manipulation of sleep is of obvious importance in illnesses such as insomnia, narcolepsy, and sleep apnea (247, 248). In addition, it may be valuable in the treatment of such conditions as psychosis and depression, where sleep disturbances are an important component of the illness. For example, delirium tremens might be best understood as a psychotic episode which is the result of an aspect of sleep emerging into wakefulness. The range and breadth of both the basic questions and the potential application of sleep research portend an exciting future for this field.

ACKNOWLEDGMENTS

We are deeply indebted to Sue Poage and Rosemary Schmele for their gracious and meticulous aid in the preparation of this manuscript.

We also wish to express our gratitude to Dr. William C. Dement for his comments and advice on this review, and for his colleagueship to this laboratory during the past ten years.

Literature Cited

1. Piéron, H. 1913. *Le Probleme Physiologique du Sommeil.* Paris: Masson. 520 pp.
2. Hartmann, E. 1973. *The Functions of Sleep.* New Haven: Yale Univ.
3. Hartmann, E. 1970. *Sleep and Dreaming,* ed. E. Hartmann, 308–28. Boston: Little Brown
4. Jouvet, M. 1969. *Science* 163:32–41
5. Jouvet, M. 1972. *Brain and Human Behavior,* ed. A. Karczmar, J. Eccles. Berlin: Springer
6. Jouvet, M. 1972. *Ergeb. Physiol.* 64:166–307
7. Jouvet, M. 1973. *Frontiers in Catecholamine Research,* ed. E. Usdin, S. Snyder, 751–58. Oxford: Pergamon
8. Jouvet, M. 1973. *Pharmacology and The Future of Man,* Vol. 4, *Proc. Fifth Int. Congr. Pharmacol., San Francisco,* 103–7. Basel: Karger
9. Jouvet, M. 1973. *Serotonin and Behavior,* ed. J. Barchas, E. Usdin, 385–400. New York: Academic
10. Jouvet, M. 1974. *Basic Sleep Mechanisms,* ed. O. Petrė-Quandens, J. Schlag, 207–32. New York: Academic
11. King, C. D. 1971. *Advan. Pharmacol. Chemother.* 9:1–91
12. Koella, W. P. 1969. *Biol. Psychiat.* 1:161–77
13. Koella, W. P. 1969. *Exp. Med. Surg.* 27:157–69
14. Koella, W. 1972. *The Sleeping Brain,* ed. M. Chase, 145–91. Los Angeles: Univ. Calif.
15. Koella, W. 1974. See Ref. 10, pp. 207–32
16. Laborit, H. 1972. *Psych. Med.* 4:453–86
17. Morgane, P. J., Stern, W. C. 1974. *Advan. Sleep Res.* 1:1
18. Williams, H. L., Holloway, F. A., Griffiths, W. J. 1973. *Ann. Rev. Physiol.* 24:279
19. Moruzzi, G., Magoun, H. W. 1949. *Electroencephalogr. Clin. Neurophysiol.* 1:455–73
20. Aserinsky, E., Kleitman, N. 1953. *Science* 118:273–74
21. Moruzzi, G. 1963. *The Harvey Lectures,* Series 58, 233–97. New York: Academic
22. Moruzzi, G. 1972. *Ergeb. Physiol.* 64:1–165
23. Falck, B., Hillarp, N. A., Thieme, G., Thorp, A. 1962. *J. Histochem. Cytochem.* 10:348–54
24. Barchas, J. D., Stolk, J. M., Ciaranello, R. D., Hamburg, D. A. 1971. *Psychological Assessment,* ed. P. McReynolds, 2:260–92. Palo Alto, California: Science and Behavior Books
25. Barchas, J. D., Ciaranello, R. D., Stolk, J. M., Brodie, H. K. H., Hamburg, D. A. 1972. *Hormones and Behavior,* ed. S. Levine, 235–329. New York: Academic
26. Andén, N. E., Carlsson, A., Hillarp, N. A., Magnusson, T. 1965. *Life Sci.* 4:1275–81
27. Andén, N. E. et al 1966. *Acta Physiol. Scand..* 67:313–27

28. Dahlström, A., Fuxe, K. 1964. *Acta Physiol. Scand.* 62, Suppl. 232
29. Dahlström, A., Fuxe, K. 1965. *Acta Physiol. Scand.* 64, Suppl. 247
30. Fuxe, K. 1965. *Acta Physiol. Scand.* 64:39–85 (Suppl. 247)
31. Pin, C., Jones, B., Jouvet, M. 1968. *C. R. Soc. Biol.* 162: 2136–41
32. Andén, N. E. et al 1966. *Experientia* 22:44–45
33. Andén, N. E., Fuxe, K., Hamberger, G., Hökfelt, T. 1966. *Acta Physiol. Scand.* 67:306–13
34. Dahlström, A., Fuxe, K., Olson, L., Ungerstedt, U. 1964. *Acta Physiol. Scand.* 62:485–87
35. Sourkes, T. L., Poirier, L. 1965. *Nature* 207:202–3
36. Parent, A., Poirier, L. J. 1969. *Can. J. Physiol. Pharmacol.* 47:781–87
37. York, D. H. 1970. *Brain Res.* 20: 233–51
38. Ungerstedt, U. 1971. *Acta Physiol. Scand.* 82, Suppl. 367
39. Ungerstedt, U. 1971. *6-Hydroxydopamine and Catecholamine Neurons,* ed. T. Malmfors, H. Thoenen, 101–66. Amsterdam: North-Holland
40. Shute, C. C. D., Lewis, P. R. 1967. *Brain* 90:497–520
41. Shute, C. C. D. 1969. *J. Physiol. (London)* 201:2
42. Olivier, A., Parent, A., Simard, H., Poirier, L. J. 1970. *Brain Res.* 18: 273–82
43. Jouvet, M. 1967. *Physiol. Rev.* 47:117–77
44. Jouvet, M., Renault, J. 1966. *C. R. Soc. Biol.* 160:1461–65
45. Kostowski, W., Giacalone, E., Garattini, S., Valzelli, L. 1968. *Eur. J. Pharmacol.* 4:371–77
46. Simon, R. P., Gershon, M. D., Brooks, D. C. 1973. *Brain Res.* 58:313–30
47. Morgane, P. J., Stern, W. C. 1972. *Ann. N.Y. Acad. Sci.* 193:95
48. Morgane, P. J., Stern, W. C. 1973. See Ref. 9, pp. 427–42
49. Mancia, M. 1969. *Electroencephalogr. Clin. Neurophysiol.* 27:487–503
50. Pujol, J. F., Buguet, A., Froment, J. L., Jones, B., Jouvet, M. 1971. *Brain Res.* 29:195–212.
51. Baumgarten, H. G., Björklund, A., Lachenmayer, L., Nobin, A., Stenevi, U. 1972. *Acta Physiol. Scand.* 84, Suppl. 373
52. Baumgarten, H. G. et al 1972. *J. Neurochem.* 19:1587–97
53. Jouvet, M., Pujol, J. F. 1974. *Advan. Biochem. Psychopharmacol.* 11:199–209
54. Froment, J. L., Petitjean, F., Bertrand, N., Cointy, C., Jouvet, M. 1974. *Brain Res.* 67:405
55. Sterman, M. B., Clemente, C. D. 1962. *Exp. Neurol.* 6:91–102
56. McGinty, D. J., Sterman, M. B. 1968. *Science* 160:1253–55
57. Matsumoto, J., Jouvet, M. 1964. *C. R. Soc. Biol.* 158:2137–40
58. Gottesmann, C. 1966. *C. R. Soc. Biol.* 160:2056–61
59. Tabushi, K., Himwich, H. E. 1969. *Psychopharmacologia* 16:240–52
60. Reite, M., Pegram, G. V., Stephens, L. M., Bixler, E. C., Lewis, O. L. 1969. *Psychopharmacologia* 14:12–17
61. Tissot, R. 1965. *Progr. Brain Res.* 18:175–77
62. Hartmann, E. 1966. *Psychopharmacologia* 9:242–47
63. Hoffman, J. S., Domino, E. F. 1969. *J. Pharmacol. Exp. Ther.* 170:190–98
64. Coulter, J. D., Lester, B. K., Williams, H. L. 1971. *Psychopharmacologia* 19: 134–47
65. Stern, W. C., Morgane, P. J. 1973. *Psychopharmacologia* 28:275–86
66. Brooks, D. C., Gershon, M. D., Simon, R. P. 1972. *Neuropharmacology* 11: 511–20
67. Haefely, W., Jalfre, M., Monachon, M. A. 1973. See Ref. 7, pp. 771–72
68. Haefely, W., Jalfre, M., Monachon, M. A. 1973. See Ref. 8, p. 171
69. Jalfre, M., Monachon, M. A., Haefely, W. 1972. *Proceedings of the First Canadian International Symposium on Sleep,* ed. D. J. McClure, 155–85. Montreal: Roche Sci. Serv.
70. Delorme, F., Froment, J. L., Jouvet, M. 1966. *C. R. Soc. Biol.* 160:2347–51
71. Sanders-Bush, E. 1973. See Ref. 9, pp. 191–200
72. Koe, B. K., Weissman, A. 1966. *J. Pharmacol. Exp. Ther.* 154:499–516
73. Weissman, A. 1973. See Ref. 9, pp. 235–48
74. Deguchi, T., Barchas, J. D. 1972. *Nature New Biol.* 235:92–93
75. Deguchi, T., Barchas, J. D. 1972. *Molec. Pharmacol.* 8:770–79
76. Deguchi, T., Sinha, A. K., Barchas, J. D. 1973. *J. Neurochem.* 20:1329–36
77. Ursin, R. 1972. *Acta Physiol. Scand.* 86:278
78. Dement, W. C. 1969. *Sleep, Physiology and Pathology, A Symposium,* ed A. Kales, 245. Philadelphia: Lippincott

79. Mouret, J. R., Bobillier, P., Jouvet, M. 1968. *Eur. J. Pharmacol.* 5:17–22
80. Rechtschaffen, A., Lovell, R. A., Freedman, D. X., Whitehead, P. K., Aldrich, M. 1969. *Psychophysiology* 6:223
81. Rechtschaffen, A., Lovell, R. A., Freedman, D. X., Whitehead, P. K. Aldrich, M. 1973. See Ref. 9, pp. 401–18
82. Cohen, H. B. et al 1970. *Proc. 78th Ann. Convention Am. Psych. Assoc.*, 831–32
83. Dement, W. C., Mitler, M. M., Henriksen, S. J. 1972. *Rev. Can. Biol.* 31:239–46
84. Wyatt, R. J. 1972. *Biol. Psychiat.* 5:33–64
85. Chernik, D. A., Ramsey, T. A., Mendels, J. 1973. *Brit. J. Psychiat.* 122: 191–99
86. Henriksen, S., Dement, W., Barchas, J. 1974. *Advan. Biochem. Psychopharmacol.* 11:169
87. Dement, W. C., Henriksen, S., Jacobs, B., Mitler, M. M. 1973. See Ref. 8, pp. 74–89
88. Saavedra, J. M., Axelrod, J. 1974. *Advan. Biochem. Psychopharmacol.* 10: 135–39
89. Stolk, J., Barchas, J., Dement, W., Schanberg, S. 1969. *Pharmacologist* 11:258
90. Cohen, H. B., Dement, W. C., Barchas, J. D. 1973. *Brain Res.* 53:363
91. Dement, W. C., Henriksen, S. J., Ferguson, J. 1973. See Ref. 9, pp. 419–24
92. Aghajanian, G. K., Rosecrans, J. A., Sheard, M. H. 1967. *Science* 156:402–3
93. Kostowski, W., Giacalone, E., Garattini, S., Valzelli, L. 1969. *Eur. J. Pharmacol.* 7:170–76
94. Padjen, A., Randic, M. 1970. *Brit. J. Pharmacol.* 39:1
95. Gumulka, W., Samanin, R., Garattini, S., Valzelli, L. 1969. *Eur. J. Pharmacol.* 8:380–84
96. Gumulka, W., Samanin, R., Valzelli, L., Consolo, S. 1971. *J. Neurochem.* 18:533–34
97. Kostowski, W. 1971. *Brain Res.* 31: 151–57
98. Polc, P., Monnier, M. 1970. *Brain Res.* 22:47–63
99. Jacobs, B. L., Asher, R., Dement, W. C. 1973. *Physiol. Behav.* 11:489–95
100. Radulovacki, M., Karmos, G. 1974. *Brain Res.* 67:542–45
101. Rechtschaffen, A., Ledecky-Janecek, S., Lovell, R. 1968. *Psychophysiology* 5:211
102. Morgane, P. J., Stern, W. C. 1973. *Brain Res.* 50:205
103. Kupfer, D. J., Foster, F. G., Bowers, M. 1972. *Brain Res.* 41:205
104. Hartmann, E. 1968. *Psychophysiology* 4:390
105. Hartmann, E., Bridwell, T. J. 1971. *Psychophysiology* 7:313
106. Monnier, M., Tissot, R. 1958. *Helv. Physiol. Pharmacol. Acta* 16:255–67
107. Macchitelli, F. J., Fischetti, D., Montarelli, N. 1966. *Psychopharmacologia* 9:447–56
108. Costa, E., Pschedit, G. R., Van Meter, W. G., Himwich, H. E. 1960. *J. Pharmacol. Exp. Ther.* 130:81–88
109. Bogdanski, D. F., Weissbach, H., Udenfriend, S. 1958. *J. Pharmacol. Exp. Ther.* 122:182–94
110. Oswald, I. 1968. *Pharmacol. Rev.* 20:273
111. Hartmann, E. 1970. *Psychophysiology* 7:320–21
112. Wyatt, R. J. et al 1971. *Electroencephalogr. Clin. Neurophysiol.* 30:505–9
113. Jouvet, M., Vimont, P., Delorme, F. 1965. *C. R. Soc. Biol.* 159:1595–99
114. Mouret, J. R., Vilppula, A., Frachon, N., Jouvet, M. 1968. *C. R. Soc. Biol.* 162:914–17
115. Aghajanian, G. K., Graham, A. W., Sheard, M. H. 1970. *Science* 169:1100
116. McGinty, D. J., Harper, R. M., Fairbanks, M. K. 1973. See Ref 9, pp. 267–80
117. Bloom, F. E., Hoffer, B. J., Nelson, C. N., Sheu, Y., Siggins, G. R. 1973. See Ref. 9, pp. 249–62
118. Sinha, A. K., Henriksen, S. J., Dement, W. C., Barchas, J. D. 1973. *Am. J. Physiol.* 224:381–83
119. Mitler, M. M. et al 1973. *Pharmacol. Biochem. Behav.* 1:501–7
120. Wyatt, R. J., Neff, N. H., Vaughan, T., Franz, J., Ommaya, A. 1974. *Advan. Biochem. Psychopharmacol.* 11:193
121. Radulovacki, M. 1973. *Brain Res.* 50:484–88
122. Weiss, E. et al 1968. *Psychophysiology* 5:209
123. Hery, F., Pujol, J. F., Lopez, M., Macon, J., Glowinski, J. 1970. *Brain Res.* 21:391–403
124. Sinha, A. K., Deguchi, T., Ciaranello, R., Dement, W. C., Barchas, J. D. 1973. *J. Neurochem.* 20:1289
125. Cramer, H., Tagliamonte, A., Tagliamonte, P., Perez-Cruet, J., Gessa, G. L. 1973. *Brain Res.* 54:372–75
126. Jones, B. E., Bobillier, P., Pin, C., Jouvet, M. 1973. *Brain Res.* 58:157
127. Petitjean, F., Jouvet, M. 1970. *C. R. Soc. Biol.* 164:2288–92
128. Buguet, A., Petitjean, F., Jouvet, M. 1970. *C. R. Soc. Biol.* 164:2293–98

129. Henley, K., Morrison, A. 1969. *Psychophysiology* 6:245
130. Jones, B. E., Harper, S., Halaris, A. In preparation
131. Hartmann, E., Chung, R., Draskoczy, P. R., Schildkraut, J. J. 1971. *Nature* 233:425–27
132. Lidbrink, P., Fuxe, K. 1973. *J. Pharm. Pharmacol.* 25:84–86
133. Lidbrink, P. 1974. *Brain Res.* 74:19
134. Petitjean, F., Laguzzi, R., Sordet, F., Jouvet, M., Pujol, J. F. 1972. *Brain Res.* 48:281–93
135. Laguzzi, R., Petitjean, F., Pujol, J. F., Jouvet, M. 1972. *Brain Res.* 48:295–310
136. Panksepp, J., Jalowiec, J. E., Morgane, P. J., Zolovich, A. J., Stern, W. C. 1973. *Exp. Neurol.* 41:233
137. Zolovich, A. J., Stern, W. C., Jalowiec, J. E., Panksepp, J., Morgane, P. J. 1973. *Pharmacol. Biochem. Behav.* 1:557
138. Stein, D., Jouvet, M., Pujol, J. F. 1974. *Brain Res.* 72:360
139. Fujimori, M., Himwich, H. E. 1971. *Fed. Proc.* Abstr. 541
140. Weitzman, E. D., McGregor, P., Moore, C., Jacoby, J. 1969. *Life Sci.* 8:751–58
141. Marantz, R., Rechtschaffen, A. 1967 *Percept. Motor Skills* 25:805–8
142. Torda, C. 1968. *Brain Res.* 10:200
143. Hartmann, E., Bridwell, T. J., Schildkraut, J. J. 1971. *Psychopharmacologia* 21:157–64
144. Marantz, R., Lovell, R. A., Whitehead, P. K. 1968. *Commun. Behav. Biol.* 2:161–64
145. King, D., Jewett, R. E. 1971. *J. Pharmacol. Exp. Ther.* 177:188–94
146. Henriksen, S., Jacobs, B., Dement, W., Barchas, J. 1973. See Ref. 7, pp. 759–62
147. Stern, W. C., Morgane, P. J. 1973. *Biol. Psychiat.* 6:301
148. Wyatt, R. J., Fram, D. H., Kupfer, D. J., Snyder, F. 1971. *Arch. Gen. Psychiat.* 24:145–55
149. Vaughan, T., Wyatt, R. J., Green, R. 1972. *Psychophysiology* 9:96
150. Dusan-Peyrethon, D., Froment, J. L. 1968. *C. R. Soc. Biol.* 162:2141–45
151. Lidbrink, P., Corrodi, H., Fuxe, K., Olson, L. 1973. See Ref. 8, pp. 90–102
152. Baekeland, F., Lundwall, L. 1971. *Electroencephalogr. Clin. Neurophysiol.* 31:284
153. Matsumoto, J., Watanabe, S. 1967. *Proc. Jap. Acad.* 43:680–83
154. Dunleavy, D. L. F., MacLean, A. W., Oswald, I. 1971. *Psychopharmacologia* 21:101–10
155. Wyatt, R. J., Gillin, J. C. *Int. J. Psychiat.* In press
156. Delorme, F., Jeannerod, M., Jouvet, M. 1965. *C. R. Soc. Biol.* 159:900–3
157. Seiden, L. S., Hanson, L. C. F. 1964. *Psychopharmacologia* 6:239–45
158. Seiden, L. S., Peterson, D. D. 1968. *J. Pharmacol. Exp. Ther.* 163:84–90
159. Stolk, J. M., Rech, R. H. 1970. *Neuropharmacology* 9:249–65
160. Janssen, P. A. J., Neimegeers, C. J. E., Schellekens, K. H. L., Lenaerts, F. M. 1967. *Arz. Forsch.* 17:841–54
161. Rechtschaffen, A., Maron, L. 1964. *Electroencephalogr. Clin. Neurophysiol.* 16:438–45
162. Baekeland, F. 1967. *Psychopharmacologia* 11:388–96
163. Hishikawa, Y., Nakai, K., Hidenobu, I., Kaneko, Z. 1965. *Electroencephalogr. Clin. Neurophysiol.* 19:518–21
164. Hartmann, E. 1969. *Biol. Psychiat.* 1:243–58
165. Zung, W. W. 1969. *Exp. Med. Surg.* 27:124–37
166. Holman, R. B., Shillito, E. E., Vogt, M. 1971. *Brit. J. Pharmacol.* 43:685–95
167. Stein, L., Wise, C. D. 1969. *J. Comp. Physiol. Psychol.* 67:189–98
168. Schaeppi, U., Koella, W. P. 1966. *Fed. Proc.* 25:628
169. Bartholini, G., Birlard, W. P., Pletscher, A., Bates, H. M. 1967. *Nature* 215:852–53
170. Kadzielawa, K., Widy-Tyszkiewicz, E. 1970. *Electroencephalogr. Clin. Neurophysiol.* 28:259–66
171. Reis, D. J., Moorhead, D. T., Meslino, N. 1970. *Arch. Neurol. (Chicago)* 22: 31–39
172. Wyatt, R. J., Chase, T. N., Scott, J., Snyder, F., Engelman, K. 1970. *Nature* 228:999–1001
173. Gillin, J. C. et al 1973. *Electroencephalogr. Clin. Neurophysiol.* 35:181–86
174. Ng, K. Y., Chase, T. N., Colburn, R. W., Kopin, I. J. 1970. *Science* 170:76–77
175. Havlicek, V. 1967. *Int. J. Neuropharmacol.* 6:83–89
176. Gillin, J. C. et al 1972. *Sleep Res.* 1:45
177. Chu, N., Bloom, F. E. 1973. *Science* 179:908–10
178. Bubenick, G., Monnier, M. 1972. *Exp. Neurol.* 35:1–12
179. Pujol, J. F., Mouret, J., Jouvet, M., Glowinski, J. 1968. *Science* 159:112–14
180. Mantegazzini, P., Pepeu, G. 1964. *J. Physiol. (London)* 173:20
181. Pepeu, G., Mantegazzini, P. 1964. *Science* 145:1069–70

182. Celesia, G. G., Jasper, H. H. 1966. *Neurology* 16:1053–64
183. Beani, L., Bianchi, C., Santinoceto, L., Marchetti, P. 1968. *Int. J. Neuropharmacol.* 7:469–81
184. Domino, E. F., Stawiski, M. 1971. *Psychophysiology* 7:315–16
185. Wikler, A. 1952. *Proc. Soc. Exp. Biol. (N.Y.)* 79:261–65
186. Jouvet, M. 1962. *Arch. Ital. Biol.* 100:125–206
187. Sagales, T., Erill, S., Domino, E. F. 1969. *Clin. Pharmacol. Ther.* 10: 522–29
188. Jacobs, B. L., Henriksen, S. J., Dement, W. C. 1972. *Brain Res.* 48:406–11
189. Vimont-Vicary, P., Jouvet-Mounier, D., Delorme, F. 1966. *Electroencephalogr. Clin. Neurophysiol.* 20:439–49
190. Henriksen, S. J., Jacobs, B. L., Dement, W. C. 1972. *Brain Res.* 48:412–16
191. Collier, B., Mitchell, J. F. 1966. *Nature* 210:424
192. Bartolini, A., Pepeu, G. 1967. *Brit. J. Pharmacol. Chemother.* 31:66–74
193. Collier, B., Mitchell, J. F. 1967. *J. Physiol. (London)* 188:83–99
194. Szerb, J. C. 1967. *J. Physiol. (London)* 192:329–45
195. Pepeu, G., Bartolini, A. 1967. *Boll. Soc. Ital. Biol. Sper.* 43:1409–11
196. Bremer, F., Chatannot, J. 1949. *Arch. Int. Physiol.* 57:106–9
197. Domino, E. F., Yamamoto, K. 1965. *Science* 150:637–38
198. Yamamoto, K. I., Domino, E. F. 1965. *Int. J. Neuropharmacol.* 4:359–73
199. Yamamoto, K. I., Domino, E. F. 1967. *Int. J. Neuropharmacol.* 6:357–75
200. Kawamura, H., Domino, E. F. 1969. *Int. J. Neuropharmacol.* 8:105–17
201. Villablanca, J. 1966. *Arch. Biol. Med. Exp.* 3:118–29
202. Cordeau, J. P., Moreau, A., Beaulnes, A., Lawrin, C. 1963. *Arch. Ital. Biol.* 101:30–47
203. Hernández-Péon, R. 1965. *Sleep Mechanisms,* ed. K. Akert, C. Bally, J. P. Schadé, 96–117. Amsterdam: Elsevier
204. Hernández-Péon, R., Chavez-Ibarra, G. 1963. *Electroencephalogr. Clin. Neurophysiol.* Suppl. 24:188–98
205. Hernández-Péon, R., Chavez-Ibarra, G., Morgane, P. J., Timo-Iara, C. 1963. *Exp. Neurol.* 8:93–111
206. Hernández-Péon, R., O'Flaherty, J. J., Mazzuchelli-O'Flaherty, A. L. 1967. *Brain Res.* 4:243–67
207. Morgane, P. J. 1969. *Psychophysiology* 6:219
208. Sharpe, L. G., Otto, M. P., Bischoff, E. J., Gasinger, M. M., Othmer, E. 1973. *Sleep: Physiology, Biochemistry, Psychology, Pharmacology, Clinical Implications,* ed. W. P. Koella, P. Levin, 339–44. Basel: Karger
209. George, R., Haslett, W., Jenden, D. 1964. *Int. J. Neuropharmacol.* 3:541–52
210. Mitler, M. M., Dement, W. C. 1974. *Brain Res.* 68:335–43
211. Karczmar, A. G., Longo, U. G., Scotti de Carolis, A. 1970. *Physiol. Behav.* 5:175–82
212. Torii, S., Mitsumori, K., Inubushi, S., Yanagisawa, I. 1973. *Psychopharmacologia* 29:65–75
213. Magherini, P. C., Pompeiano, O., Thoden, U. 1971. *Brain Res.* 35:565–69
214. Bowers, M. B., Hartmann, E. L., Freedman, D. X. 1966. *Science* 153:1416–17
215. Tsuchiya, K., Toru, M., Kobayashi, T. 1969. *Life Sci.* 8:867–73
216. Jasper, H. H., Tessier, J. 1965. *Science* 147:1448–49
217. Gadea-Ciria, M., Stadler, H., Lloyd, K. G., Bartholini, G. 1973. *Nature* 243:518
218. Haranath, P. S. R. K., Venkatakrishna-Bhatt, H. 1973. *Jap. J. Pharmacol.* 23:241
219. Bessman, S. P., Fishbein, W. N. 1963. *Nature* 200:1207–8
220. Roth, R. H., Suhr, Y. 1970. *Biochem. Pharmacol.* 19:3001–12
221. Jouvet, M., Cier, A., Mounier, D., Valatx, J. L. 1961. *C. R. Soc. Biol.* 155:1313–16
222. Matsuzaki, M., Takagi, H. 1967. *Brain Res.* 4:206–22
223. Winters, W. D., Spooner, C. E. 1965. *Electroencephalogr. Clin. Neurophysiol.* 18:287–96
224. Metcalf, D. R., Ende, R. N., Stripe, J. T. 1966. *Electroencephalogr. Clin. Neurophysiol.* 20:506–12
225. Gessa, G. L. et al 1966. *Life Sci.* 5:1921–30
226. Hutchins, D. A., Rayevsky, K. S., Sharman, D. F. 1972. *Brit. J. Pharmacol.* 46:409
227. Legendre, R., Piéron, H. 1910. *C. R. Soc. Biol.* 68:1108–9
228. Schnedorf, J. F., Ivy, A. C. 1939. *Am. J. Physiol.* 125:491–505
229. Monnier, M., Koller, T., Graber, S. 1963. *Exp. Neurol.* 8:264–77
230. Monnier, M., Hösli, L. 1964. *Science* 146:796–97
231. Monnier, M., Hösli, L. 1965. *Pflügers Arch.* 282:60–75
232. Hösli, L., Monnier, M. 1965. *Pflügers Arch.* 283:17–25

233. Schneiderman, N., Monnier, M., Hösli, L. 1966. *Pflügers Arch.* 288:65–80
234. Monnier, M., Hatt, A. M. 1971. *Pflügers Arch.* 329:231–43
235. Schoenenberger, G. A., Cueni, L. B., Monnier, M., Hatt, A. M. 1972. *Pflügers Arch.* 338:1–17
236. Monnier, M., Dudler, L., Schoenenberger, G. A. 1973. *Pflügers Arch.* 345:23–35
237. Schoenenberger, G. A., Cueni, L. B., Hatt, A. M., Monnier, M. 1973. See Ref. 208, pp. 263–66
238. Ringle, D. A., Herndon, B. L. 1968. *Pflügers Arch.* 303:344–49
239. Ringle, D. A., Herndon, B. L. 1969. *Pflügers Arch.* 306:320–28
240. Pappenheimer, J. R., Miller, T. B., Goodrich, C. A. 1967. *Proc. Nat. Acad. Sci. USA* 58: 513–17
241. Fencl, V., Koski, G., Pappenheimer, J. R. 1971. *J. Physiol. (London)* 216: 565–89
242. Haulica, I., Ababei, L., Branisteanie, D., Topoliceanu, F. 1973. *J. Neurochem.* 21:1019–20
243. Drucker-Colin, R. R., Rojas-Ramirez, J. A., Vera-Treuba, J., Monroy-Ayala, G., Hernández-Péon, R. 1970. *Brain Res.* 23:269–73
244. Drucker-Colin, R. R. 1973. *Brain Res.* 56:123–34
245. Kessler, S., Ciaranello, R. D., Shire, J. G. M., Barchas, J. D. 1972. *Proc. Nat. Acad. Sci. USA* 69:2448–50
246. Stolk, J. M., Conner, R. L., Levine, S., Barchas, J. D. 1974. *J. Pharmacol. Exp. Ther.* 90:193–209
247. Guilleminault, C., Elridge, F., Dement, W. 1973. *Science* 181:856–58
248. Dement, W., Guilleminault, C. 1973. *Hospital Pract.* 8:57–71

Copyright 1975. All rights reserved

HYPERGLYCEMIA, POLYOL METABOLISM, AND COMPLICATIONS OF DIABETES MELLITUS

❖7170

Kenneth H. Gabbay, M.D.[1]
Department of Pediatrics, Children's Hospital Medical Center and Harvard Medical School, Boston, Massachusetts 02115

INTRODUCTION

The discovery of insulin and its therapeutic use in diabetes has changed the natural history of this disease. Today, the longevity afforded the patient by insulin is associated with the indiscriminate and sometimes crippling long term development of retinopathy, neuropathy, nephropathy, cataracts, and generalized atherosclerosis (1, 2). It is becoming increasingly apparent that these manifestations of diabetes are an intrinsic part of basic and as yet unknown disease processes which affect the whole organism. The processes leading to the chronic damage in diabetes may be related to subtle alterations in intracellular glucose metabolism in diabetic cells, with or without attendant hyperglycemia. The tissues bearing the brunt of diabetic manifestations (lens, nerve, retina, kidney, blood vessels, and islet cells) are freely permeable to glucose and do not require insulin for glucose penetration as do muscle and adipose tissues, and hence are exposed to the ambient blood glucose levels. The relationship of the blood sugar control to the long term development of diabetic manifestations is problematic. There is definite evidence of a beneficial effect of "good" control of blood glucose in reducing the incidence of infections and diabetic ketoacidosis (3, 4), as well as evidence of a short term beneficial effect on some other diabetic manifestations (5). However, a consensus on a convincing long term beneficial relationship to the prevention of diabetic complications is still lacking, a situation which may well be due to our inability to achieve optimal blood glucose control with current therapeutic regimens (6).

---

[1] Supported by U.S. Public Health Service Grant #AM-15019, a grant from the National Foundation—March of Dimes, and an Established Investigatorship of the American Heart Association.

Recent investigations of accessory pathways of glucose metabolism in tissues indicate a role for some pathways in the formation of some diabetic complications. This chapter will review current knowledge of the sorbitol pathway and its possible significance in diabetes and galactosemia.

THE SORBITOL PATHWAY

General

In 1924 Orr described the presence of fructose in human fetal blood (7). Hubbard & Russell (8) reported the finding of free fructose in human cerebrospinal fluid. Mann (9) subsequently characterized the main sugar of seminal fluid as fructose and demonstrated in rabbits that fructose is derived from blood glucose and is responsive to fluctuations of the latter in diabetes with and without insulin therapy (10). He further showed (11) that seminal fluid fructose content is directly influenced by hormonal factors, virtually disappearing in castrated and hypophysectomized animals. Hers (12) later demonstrated that seminal fructose is formed in the seminal vesicles from blood glucose, primarily through the activity of the "sorbitol pathway." This pathway consists of two enzymes that are able to convert glucose to fructose via the intermediate step of sorbitol.

$$\text{D-Glucose} + \text{NADPH} + \text{H}^+ \xrightarrow[\text{Reductase}]{\text{Aldose}} \text{Sorbitol} + \text{NADP}^+$$

$$\text{Sorbitol} + \text{NAD}^+ \underset{\text{Dehydrogenase}}{\overset{\text{Sorbitol}}{\rightleftarrows}} \text{D-Fructose} + \text{NADH} + \text{H}^+$$

Later work by Hers (13) showed that the pathway is responsible for the formation of sorbitol in sheep placenta, and further conversion to fructose by the fetal liver. Fructose is the main fetal blood sugar in sheep, disappearing within hours after the separation of the fetus from the placenta (14).

ALDOSE REDUCTASE Aldose reductase catalyzes the conversion of free glucose to its sugar alcohol sorbitol. It possesses broad substrate specificity for many aldoses, and is characterized by low affinity for glucose and galactose (15–21). The low affinity for hexoses, and the availability of excess amounts of free hexoses in diabetes and galactosemia cause increased formation of the sugar alcohols sorbitol and galactitol respectively. Aldose reductase requires NADPH for its activity, and since NADPH is provided in the cell primarily by the action of the hexose monophosphate shunt (HMP shunt), a relationship between the HMP shunt and the sorbitol pathway has been described in lens and peripheral nerve (22, 23). In vitro, aldose reductase activity is enhanced in the presence of mercaptoethanol and stimulated by lithium sulfate. The enzyme displays allosteric properties and exists in an equilibrium between an active form and two inactive subunits (24). The coenzyme product of its reaction, NADP$^+$, inhibits the enzyme at physiological concentrations by altering the equilibrium in favor of the inactive subunits. The reaction catalyzed by aldose reductase is virtually unidirectional, the equilibrium favoring the forward reaction.

An enzyme of the glucuronic acid-xylulose shunt, NADP-L-hexonate dehydrogenase, which has many similarities to aldose reductase, is also present in many tissues. Hexonate dehydrogenase (16–19, 25, 26) has a poor ability to convert hexoses to their respective sugar alcohols (K_m 0.5–2 M), and is most active in the reduction of uronic acids. Since this enzyme is present in many tissues, care must be taken to differentiate between it and aldose reductase. The two enzymes may be differentiated by different mobilities on DEAE-cellulose column chromatography, reactivity with glucuronate and glyceraldehyde as substrates (a ratio <1 for aldose reductase), and immunologically (19).

SORBITOL DEHYDROGENASE The further conversion of sorbitol to fructose is catalyzed by sorbitol dehydrogenase. This ubiquitous enzyme has broad substrate specificities for many sugar alcohols, converting them to the respective keto sugars, i.e. sorbitol and mannitol to fructose, xylitol to D-xylulose, and ribitol to D-ribulose, etc. Galactitol and D-arabitol are poor substrates for this enzyme. This latter property is of significance in galactosemia where the inability to further metabolize galactitol leads to enhanced accumulations of the latter in galactosemic tissues. Sorbitol dehydrogenase has been isolated and purified from a number of tissues (27–31) and has been shown to be hormonally sensitive during the development of the accessory sexual organs (32). The reaction catalyzed by sorbitol dehydrogenase is reversible, and NADPH can be utilized in the reverse action as well as NADH. The forward reaction utilizes NAD^+ only. The human and rat liver and kidney enzymes differ from the sheep liver enzyme in being able to use mannitol as a substrate. The K_m mannitol for the human enzyme is 100 mM, and for the rat enzyme 5–10 mM. The enzyme from all sources can utilize sorbitol with a lower K_m of 0.5 mM (31).

Tissue Distribution

The presence of sorbitol and fructose in substantial amounts in tissues is prima facie evidence for the presence of the sorbitol pathway. Enzymatic determination of sorbitol by sorbitol dehydrogenase is complicated by the broad substrate specificity for other sugar alcohols displayed by this enzyme. Therefore, rigorous proof of the identity of sorbitol by gas liquid chromatography of the trimethylsilyl and acetate derivatives is necessary. For example, by use of gas chromatography and mass spectroscopy, we identified the gram quantities of sugar alcohol excreted in human diabetic urines as mannitol rather than sorbitol (33). The best criterion for establishing the presence of the sorbitol pathway in tissues is the identification of aldose reductase. Again, care to distinguish this enzyme activity from that of hexonate dehydrogenase is required since the latter enzyme is present in many tissues. Aldose reductase identification in milligram amounts of tissues is now possible by immunological techniques using a specific antibody to aldose reductase which does not interact with hexonate dehydrogenase (19). Figure 1 demonstrates the presence of aldose reductase by the Ouchterlony technique in a human insulinoma using a monospecific antibody to human placental aldose reductase.

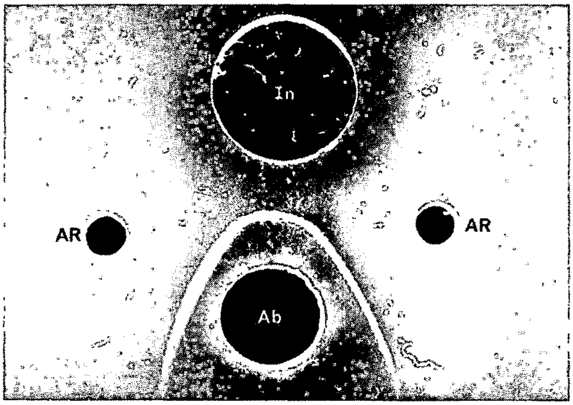


Figure 1 Ouchterlony double immunodiffusion plate demonstrating the presence of aldose reductase in a crude homogenate of a human insulinoma (In) (Equal to approximately 10 mg tissue), and its identity with purified human placental aldose reductase (AR). The bottom well contains rabbit anti-human placental aldose reductase serum.

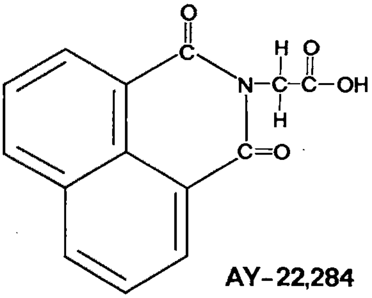
Aldose reductase appears to be highly localized to certain cell types within tissues. For instance, the highest concentration of aldose reductase is present in the lens epithelium (34), while it is entirely localized to the schwann cell in peripheral nerve (35), to the kidney papilla (36), and to the islets of Langerhans in pancreas (17 and Figure 1). These localizations are important for correlating the site of sugar alcohol formation with the site of pathology, and additionally imply much higher local concentrations of sugar alcohol than would be apparent from determining whole tissue levels. Table 1 presents a list of tissues in which the presence of aldose reductase, hexonate dehydrogenase, and sorbitol dehydrogenase has been clearly established. Additional specific cellular and subcellular localizations of these enzymes are necessary to further clarify the normal role of this enzymatic pathway.

Aldose Reductase Inhibitors

A number of specific inhibitors of aldose reductase have been described (17, 24, 37). Glutaric acid and a number of its analogues such as 3, 3-tetramethylene glutaric acid (TMG), and dimethylglutaric acid (DMG) have been very useful in studies of the role of the sorbitol pathway in tissues in vitro (17, 24). While the glutaric acid analogues are useful in in vitro studies, they proved ineffective in in vivo studies because of absorption and penetration limitations. Colchicine was also found to be a potent aldose reductase inhibitor (17), and more recently an orally effective inhibitor of aldose reductase, 1, 3-dioxo-1H-benz[de]isoquinoline-2(3H) acetic acid (AY-22, 284), was described (38), which proved quite potent in experimental animals with virtually no side effects or toxicity at dose levels far exceeding therapeutic levels (Figure 2). These inhibitors are important in clarifying the relationship of sorbitol accumulation to the formation of diabetic and galactosemic manifestations, and may be of possible future clinical use.

Table 1 Tissue distribution of aldose reductase and related enzymes

Tissue	Aldose Reductase (AR)	Hexonate Dehydrogenase (HDH)	Sorbitol Dehydrogenase (SDH)	Localization
Adipose	+	—	—	
Adrenal	+	0	+	Unknown
Aorta	+	+	+	Unknown
Brain	+	+	+	Unknown
Kidney				
Cortex	0	+	+	AR in papilla; HDH In cortex; SDH mostly In cortex (@ 90%); AR absent in glomeruli
Medulla	±	+	+	
Papilla	+	±	+	
Lens	+	0	+	AR highest in lens epithelium
Liver	+	+	+	Unknown
Muscle	+	—	—	Unknown
Pancreas	+	+	+	AR + SDH in islets of Langerhans; HDH localization unknown
Placenta	+	0	0	Unknown
Red blood cells	0	+	+	AR absent
Seminal vesicles	+	—	+	Unknown
Spinal cord	?	+	+	Unknown
Spinal roots and peripheral nerves	+	0	+	AR in Schwann cell; SDH primarily in axon (@ 90%)
Testes	0	—	+	Unknown



AY-22,284

1,3-dioxo-1H-benz[de]isoquinoline-2(3H) acetic acid

Figure 2 The chemical structure of the aldose reductase inhibitor AY-22, 284.

ROLE OF SORBITOL PATHWAY IN DIABETIC AND GALACTOSEMIC MANIFESTATIONS

The high K_m aldose reductase pathway of metabolism of free glucose and galactose appears to be relatively dormant in normal tissues. In diabetes and galactosemia, the availability of a large pool of free intracellular aldoses causes increased formation of the sugar alcohols sorbitol and galactitol respectively. Once formed, the sugar alcohols are trapped intracellularly because of their inability to penetrate cell membranes efficiently (39, 40). Sorbitol can be converted to fructose, but since fructose is poorly metabolized in some of the tissues, the net effect is an accumulation of both sorbitol and fructose intracellularly with resultant hypertonicity and osmotic consequences. The inability of the sorbitol dehydrogenase enzyme to further metabolize galactitol results in even greater accumulations of this sugar alcohol in galactosemic tissues. The intracellular hypertonicity causes osmotic swelling in tissues because the accumulation of 1 μmole of solute per gram of tissue water is equivalent to 1 mosmole per liter in osmotic pressure. Although the sorbitol pathway has been demonstrated in many tissues (Table 1), only three-lens, sciatic nerve, and renal papilla have thus far been shown to contain sufficiently high concentrations of sugar alcohol in diabetes and galactosemia to be of possible osmotic importance. Whether the osmotic mechanism described is valid for other tissues as well is uncertain. Obviously, if the aldose reductase enzyme is localized to specific cell types, the local accumulations may indeed be osmotically significant. At the present time there is no convincing evidence to indicate that the sugar alcohol per se has any deleterious effect on the tissues by any mechanism other than the osmotic effect mentioned above.

Mechanism of Cartaract Formation

Osmotic effects in human diabetic cataracts have long been suspected. Myopia as a diabetic symptom was known as early as 1859 (41), and Duke-Elder (42) concluded that in hyperglycemic patients, myopia occurs owing to the swelling of the lens, whereas hypermetropia results when the blood sugar is lowered. The histological picture seen in diabetic and galactosemic cataract can be reproduced in experimental diabetic and galactose-fed rats. The first sign is the appearance of hydropic lens fibers caused by the accumulation of fluid and in the more advanced stages the swelling continues until the fibers liquify and rupture leading to intrafibrillar clefts. An explanation for the lenticular swelling became possible when Van Heyningen (43) found increased levels of sorbitol, galactitol, and xylitol in diabetic, galactose-induced, and xylose-induced cataractous rat lenses, respectively.

The availability of high concentrations of these sugars in the aqueous humor and the ability of these aldoses to penetrate into the lens leads to excess intracellular concentrations of sugars that are then converted to the respective sugar alcohols by the high K_m pathway of aldose reductase. Kinoshita et al (44–49) demonstrated that these accumulations of sugar alcohols were paralleled by the accumulation of water and lenticular swelling. It was subsequently demonstrated that an influx of sodium ions accompanies the water swelling, intracellular potassium is lost to the medium, and the amino acid concentrating mechanism stops functioning. Losses of ATP, reduced glutathione, small peptides, and proteins occur. Free intracellular myoinositol is lost to the medium in proportion to the sugar alcohol acculations. This process of swelling and electrolyte imbalance eventually results in a precipitous loss of osmotic integrity, with massive influx of sodium and chloride ions and water and the sudden appearance of the lenticular opacity. The lenticular opacity represents the macroaggregation or precipitation of the normally translucent lenticular proteins. This sequence of events, which is summarized in Figure 3, occurs at a more rapid pace in galactose cataracts or in lenses incubated in tissue culture at high galactose levels. The larger galactitol accumulations in the galactosemic model exert greater osmotic and electrolyte imbalances and serve as a magnifying model in the study of the effects of sugar alcohol accumulations. The difference in magnitude between galactitol and sorbitol accumulations (80 μmoles vs 30 μmoles/g of lens respectively) is reflected in the much earlier and more predictable cataract formation in rats fed galactose than in diabetic rats (16 days vs 3 months).

The validity of the role of sugar alcohol accumulation and the consequent osmotic effect in the mechanism of cataract formation was demonstrated with the use of the aldose reductase inhibitor TMG. This inhibitor prevented the galactitol and sorbitol formation in lenses incubated in media containing high galactose and high glucose levels, respectively (37, 49). All secondary changes such as osmotic swelling, vacuole formation, amino acid losses, electrolyte imbalances, and the opacity formation were also prevented. More recently, AY-22, 284 successfully prevented galactosemic cataract formation in experimental rats by blocking the sugar alcohol accumulations and their consequent effects (38, 50). These in vitro and in vivo experiments clarified the pathogenesis of diabetic and galactosemic cataracts, and demonstrated that inhibition of aldose reductase affords a possible means of preventing this complication in diabetes and galactosemia (Figure 4).

While diabetic cataracts are rather infrequent in juvenile diabetic patients, extensive changes are seen in their lenses that fall short of final opacity formation. These changes are appreciated clinically as swelling of the lenses with consequent blurring of vision and changing refractive index occurring during periods of persistent hyper-

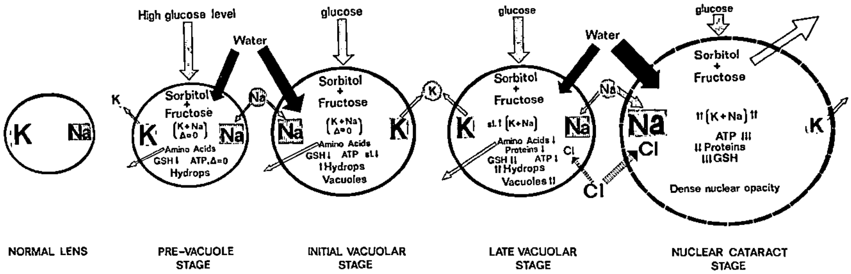


Figure 3 Scheme of changes occurring during the development of a sugar cataract (modified from Kinishita, J. H., 34). Sizes of arrows reflect relative magnitudes of alterations. Lenticular sizes indicate swelling (reproduced from Gabbay, K. H. 1973. *N. Engl. J. Med.* 288:831).

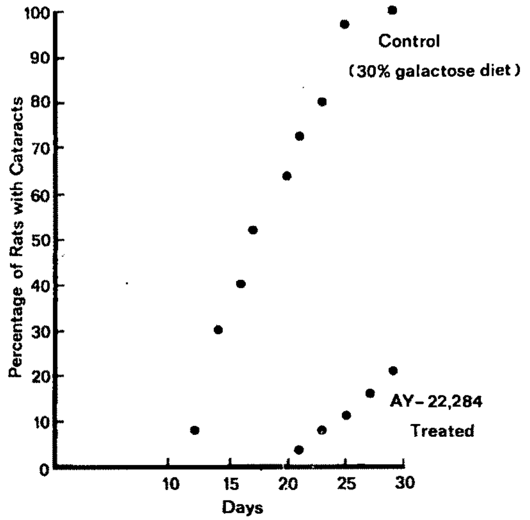


Figure 4 Effect of AY-22, 284 (0.7% of diet or approximately 0.96 g/kg per day) on formation of macroscopically detectable cataracts in rats fed a diet containing 30% galactose (from Dvornik, D. et al, 38).

glycemia. Though these early changes are reversible with return of blood glucose to normal, a large percentage of the lenses of these juvenile diabetic patients will show small opaque spots on the anterior lens surface on slit lamp microscopy. These small opacities do not interfere with vision, as opposed to the dense nuclear cataracts. It should be emphasized that the senile cataracts seen in older nondiabetic and diabetic patients differ morphologically and metabolically from the metabolic cataracts described above. However it is possible that sorbitol accumulations in older diabetic patients may act synergistically in advancing or enhancing the process of senile cataract formation.

Nervous System

Stewart et al (51) described the presence of free fructose in normal rabbit tibial nerves. The sorbitol pathway was subsequently demonstrated to be present in normal nerves and its activity was shown to be markedly elevated in diabetes with accumulations of sorbitol and fructose (52–55). The sum of sorbitol and fructose accumulation in these nerves is greater than 10 μmoles/g wet weight. A direct correlation between the blood glucose levels and nerve fructose levels was observed (Figure 5). In addition, diabetic sciatic nerve contained in excess of 17 μmoles of glucose/g wet weight, indicating the availability of a large free glucose compartment. In galactose-fed rats galactitol accumulations are of a greater magnitude than the sorbitol accumulations observed in diabetes (12 to 20 μmoles/g wet weight of nerve).

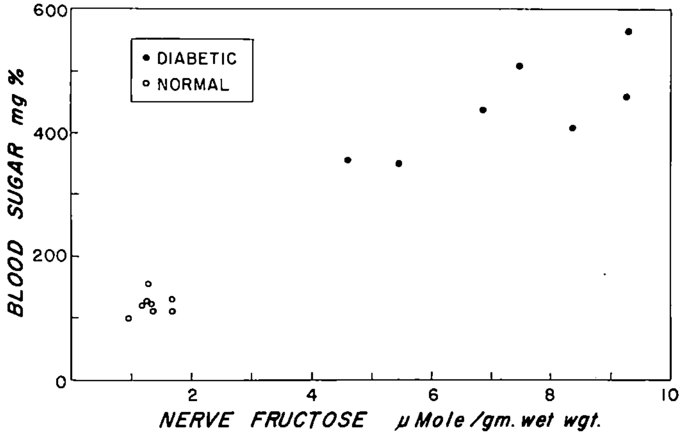


Figure 5 Relation between blood glucose levels and nerve fructose levels (from Gabbay, K. H. et al, 52).

Significant differences in the levels of sorbitol and fructose are found between the peripheral nervous system and the central nervous system, and these differences are exaggerated in the diabetic state where the total sorbitol plus fructose level in brain or spinal cord is less than 0.4 μmoles/g wet weight. Studies of the distribution of the enzymes of the sorbitol pathway in nervous tissue demonstrated that while the spinal and peripheral nerves contain aldose reductase only, the spinal cord and brain contain mainly NADP-L-hexonate dehydrogenase and a small amount of aldose reductase.. Sorbitol dehydrogenase was present in all three tissues (Table 1). Wallerian degeneration experiments indicate that the aldose reductase enzyme is localized ot the Schwann cells, which are the specific cells involved in the formation and maintenance of myelin in the peripheral nervous system (35, 56–57). Such localization provides biochemical evidence that an important metabolic disturbance such as the sorbitol accumulation is occurring at the suspected site of pathology in the diabetic nerve, namely the Schwann cell (58). These findings led to investigations of the possibility that the sorbitol pathway may play a role in the formation of diabetic neuropathy. Clinically, diabetic neuropathy is known to respond to good control of the blood glucose (5, 59), and pathologically it is characterized by segmental demyelinization associated with Schwann cell abnormalities and loss (58). Physiologically, diabetic neuropathy is characterized by decreased nerve conduction velocity in both sensory and motor nerves. In general, diabetic neuropathy appears to consist of two on-going processes, an acute reversible metabolic damage, and a more permanent irreversible damage perhaps reflecting the Schwann cell loss and segmental demyelinization. Although it is possible to demonstrate an improvement in motor nerve conduction velocity (MNCV) in newly diagnosed diabetic patients brought under blood sugar control, and deterioration when the blood sugar is

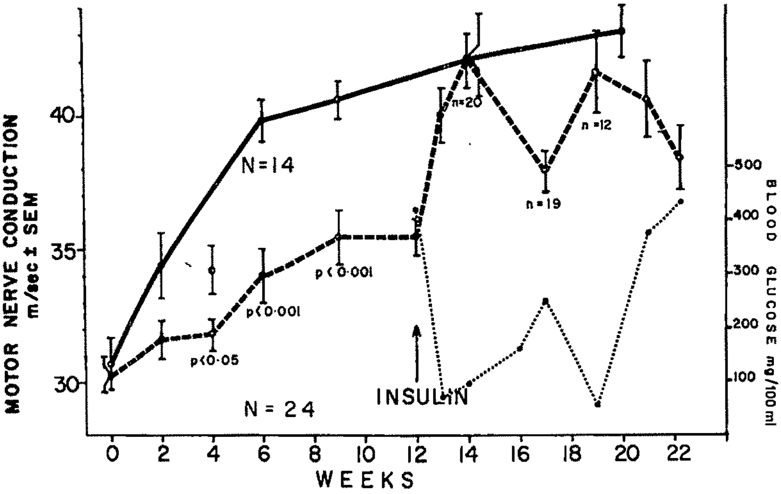


Figure 6 Serial course of motor-nerve conduction velocity in normal and streptozotocin diabetic rats over a 22 week period (reproduced from Gabbay et al, 60, with the permission of the publisher). The solid line represents normal, the dashed line diabetic, and the dotted line blood glucose.

allowed to rise (5), the long term diabetic patients have a baseline irreversible motor and sensory conduction velocity defect which is not completely responsive to blood glucose regulation. It is possible that at any one point in time we are observing a continuum or a superimposition of the two processes in the diabetic patient. For these reasons, a causal relationship of enhanced sorbitol pathway activity to diabetic neuropathy is not clear at the present time.

The MNCV defect can be reproduced in experimental diabetic rats as well as can the restoration of MNCV to normal with insulin treatment (60). Figure 6 demonstrates the induction of the MNCV defect in streptozotocin diabetic rats and the effect of regulation of the blood sugar on restoration of the MNCV to normal.[2] Nerve glucose, sorbitol, and fructose contents were reduced below normal levels in the diabetic rats kept at a hyperglycemic level, and were intermediate in the mildly hyperglycemic animals (60). Though a relationship is indicated between the sorbitol accumulations and the MNCV defect, it is not clear whether the elevated sorbitol levels are an index or a cause of neuropathy. Acute treatment of diabetic rats with AY-22, 284 reduces the sorbitol and fructose levels by 50 and 64% respectively (38), and improves the MNCV, but a definitive causal role for sorbitol accumulation in the formation of diabetic neuropathy awaits the long term prevention by aldose reductase inhibitors of permanent nerve damage despite persistent hyperglycemia.

---

[2] Serial in vivo MNCV was measured in rats acclimated for 30–60 min in a temperature controlled room (25°C).

A clear-cut role for sugar alcohol accumulations and the consequent osmotic swelling in the development of neuropathy is demonstrated in studies in galactose-fed rats. A MNCV defect is present in galactose-fed rats which is exacerbated by additional galactose gavage and remedied within two days of removal of galactose from the diet. The increase in galactitol and nerve water content closely followed the galactose intake and the MNCV defect (61). In separate experiments, AY-22, 284 successfully prevented the accumulation of galactitol and osmotic swelling in the nerves of galactose-fed rat, and the MNCV defect was also prevented or markedly delayed (62).

Although the development of the MNCV defect and its exacerbation by hyperglycemia and amelioration by insulin treatment has been demonstrated in both patients and experimental animals, a recent report by Greene et al (63) presents apparently conflicting data. These authors reported that a MNCV defect develops within 2 weeks of the induction of streptozotocin diabetes in rats which is not affected by insulin treatment. They further found that sciatic MNCV is normal in diabetic rats following 14 months of persistent untreated hyperglycemia. Since the MNCV determinations in these studies were performed with exposed rat tissue artificially heated and maintained at 35°C, it is possible that the resultant heightened conduction velocity at this temperature reflects fibers or parameters other than those normally measured in patients and experimental animals at room temperature by noninvasive techniques. Of interest, the MNCV defect in these diabetic rats was restored to normal by feeding a diet enriched with myoinositol (1%). These authors (64, 65) previously reported that feeding a diet containing 35% myoinositol causes a significant MNCV defect in otherwise normal rats. It is not yet clear whether these findings are related to or are part of the diabetic MNCV defect and neuropathy seen in patients or experimental animals.

The presence of aldose reductase and NADP-L-hexonate dehydrogenase in brain and cord was mentioned above; however, nothing is known regarding their localization. The low levels of sorbitol and fructose in these two tissues (0.05 μmoles/g and 0.08 μmoles/g respectively) are related to the low levels of aldose reductase activity and the ability of both of these tissues to utilize fructose efficiently. Significant concentrations of fructose in the cerebrospinal fluid (CSF) were first demonstrated by Hubbard & Russell (8). Wray & Winegrad (66) subsequently demonstrated that CSF fructose is elevated in diabetic patients in proportion to the elevated CSF glucose. CSF sorbitol concentrations are also elevated in diabetic patients with a high level of correlation to the blood glucose level (67). Spinowitz & Altszuler (68) demonstrated that in dogs made diabetic with growth hormone injections, which results in hyperglycemia and hyperinsulinism, CSF fructose levels and turnover rates are significantly elevated. Restoration of normoglycemia by a prolonged fast returned CSF fructose levels to normal. Short glucose infusions to elevate plasma glucose (4 hr) in normal dogs did not cause an increase of CSF fructose. Presently, it is not known whether the CSF fructose is produced entirely by the spinal nerves or whether there are any contributions from cord and brain metabolism. Clements et al (69–71) have proposed that the accumulation of sorbitol and fructose in brain is responsible for the brain edema seen following the rapid lowering of blood glucose during the treatment of nonketotic hyperglycemic coma in patients, and during the rapid lowering of blood glucose in dogs given prolonged glucose infusions. Since the levels of sorbitol and fructose in diabetic brain are virtually negligible (less than 0.4 μmoles/g or less than 0.4 mosmoles/l), it is difficult to explain the observed brain edema on this basis. A more likely possibility is a lag in brain glucose disappearance as contrasted with the rapid blood glucose decrease. Such an effect was indeed observed in a number of studies (68, 71). More recent studies (72, 73) confirm that an osmotic mechanism for the brain edema mediated by polyol pathway activity is not tenable.

Kidney

We reported the presence of the enzymes of the sorbitol pathway in rat and bovine kidneys (19, 35). As in other tissues, the distribution pattern demonstrated localization of the aldose reductase enzyme to the kidney papilla (Table 1). The cortical enzyme was characterized as NADP-L-hexonate dehydrogenase. While the main sorbitol dehydrogenase activity was localized to the cortex, approximately 5% of the total activity was present in the papilla as well. Isolated glomeruli did not have aldose reductase activity either by biochemical or immunological assay. The sugar alcohols were identified as galactitol and sorbitol in galactosemic and diabetic rats respectively. Sugar alcohols were present in appreciable levels (48 and 12 μmoles/g in galactosemic and diabetic rats respectively) in the papilla, while none could be identified in the cortex. As in lens and nerve, myoinositol content in the papilla was decreased in these two conditions. Schofield et al (74) recently found that cultured monkey kidney epithelial cells accumulate large concentrations of sugar alcohol when incubated in the presence of high concentrations of glucose or galactose (285 and 585 μmoles/g protein respectively). Sugar alcohol accumulations were suppressed during incubation with the aldose reductase inhibitor TMG. The report does not specify the type or derivation of the kidney cells.

The data available on sorbitol pathway activity in kidney are insufficient to establish whether it has a definitive role in diabetic nephropathy. The sugar alcohol accumulations in the kidney papilla are sufficiently high to be of osmotic consequence, and indeed swelling and vacuolation of the tubular cells can be observed in diabetic and galactosemic rat kidney papillae and we have suggested participation of this pathway in the interstitial nephritis and papillary atrophy seen in uncontrolled diabetic patients and experimental animals.

The Islets of Langerhans

In whole rat pancreas, approximately 40% of the total glyceraldehyde-reducing activity is NADP-L-hexonate dehydrogenase, and the remaining 60% represents aldose reductase. Aldose reductase activity is present in collagenase isolated rat islets of Langerhans while acinar fragments contain the NADP-L-hexonate dehydrogenase. More specifically, aldose reductase is present in the beta cells as demonstrated by immunological techniques using the Ouchterlony and fluorescein-conjugated antibody localization (Figure 1). The presence of glucose, sorbitol, and fructose in incubated isolated rat islets was demonstrated by Morrison et al

(75), and we have confirmed the presence of these compounds by gas chromatography.

A number of considerations suggested to us the possible involvement of the sorbitol pathway in the mechanism of insulin release. These considerations were: (*a*) the constancy of the metabolite concentrations in the glycolytic and hexose monophosphate pathways during glucose-stimulated release of insulin, (*b*) the dual requirement for NADPH of both aldose reductase and glucose mediated insulin release, (*c*) the availability of a large pool of free intracellular glucose in islets, and (*d*) the kinetic characteristics of aldose reductase. We reported (17) that glutaric acid and its analogues and colchicine are potent inhibitors of aldose reductase as well as inhibitors of both phases of glucose-induced insulin release by the perfused rat pancreas preparation. Corkey & Mayhew (76) reported that sorbitol levels in incubated islets rise within 1 min of exposure of the islets to high glucose media and prior to the initiation of insulin release. Similarly, Malaisse & Sener (77) showed that the net amount of sorbitol formed by islets over a 90-min incubation period accounted for 3 to 4% of the uptake of glucose by the islets. Aldose reductase inhibitors abolished the glucose-induced sorbitol accumulation; however, in these experiments they failed to affect glucose-induced insulin release. The mechanism and exact mode of participation of this pathway in insulin release is as yet unknown.

Other Tissues

The presence of the pathway in other tissues was described in Table 1. Morrison et al (78, 79) have proposed that the sorbitol pathway in human red blood cells plays an important role in the decrease in 2, 3-diphosphoglyceric acid (2, 3-DPG) which is observed in diabetic patients. However a number of authors (19, 26) have been unable to demonstrate the presence of aldose reductase in red blood cells, although NADPH-L-hexonate dehydrogenase was easily demonstrable. Similarly, in other studies Clements et al (80, 81) have proposed participation of the sorbitol pathway in the swelling and decreased oxygen uptake of aorta strips incubated at high glucose levels. The levels of sorbitol observed by these authors are miniscule (up to 50 nmoles/g), and cannot account via an osmotic mechanism for the swelling observed. These authors postulate some as yet unknown direct effect of sorbitol on cellular metabolism for the observed effects.

Body Fluids

The presence of sorbitol and fructose in seminal and cerebrospinal fluids was mentioned above. In addition, the aqueous humor contains an average of 5.4 mg/100 ml in normal rabbit eyes, and this level falls to 1.4 mg/100 ml in aphakic eyes. Since the blood fructose level is 1.7 mg/100 ml of plasma, it was concluded that the lens is the only ocular tissue that supplies fructose to the aqueous humor (82). Pitkanen (67) found that human serum contains both sorbitol and mannitol and that the level of both sugar alcohols were elvated in patients with chronic renal failure. Aloia (83) found that serum sorbitol is elevated in diabetic patients (0.59 mg/100 ml vs 0.24 mg/100 ml in normal serum), and that sorbitol levels were even more elevated in uremic serum (3.36 mg/100 ml). Pitkanen (84) also demonstrated that the urinary excretion of mannitol, but not sorbitol, is elevated in diabetic patients with an excretion of 90 mg of mannitol/24 hr. In a study with controlled galactosemic patients, Blau (85) found that these patients excreted a variable amount of both mannitol and galactitol which were present occasionally in gram quantities. Control subjects had no galactitol and only occasional traces of mannitol.

We reported (33) the urinary excretion of significant amounts of mannitol by diabetic patients (>3.5 gs/24 hr) which occurs randomly without correlation to the amount of glycosuria. However, the urine of uncontrolled streptozotocin diabetic rats contained only sorbitol. Because of the close relationship between fructose, sorbitol, and mannitol, and dietary differences between rats and humans, fructose was administered orally to human volunteers and rats with resultant urinary excretion of mannitol in both species. Further studies of the aldose reductase enzyme from both human and rat tissues demonstrated a previously unsuspected capacity (K_m 25 mM) to reduce fructose to a mixture of sorbitol and mannitol with both NADH and NADPH as coenzymes. Treatment of normal rats with AY-22, 284 resulted in a significant reduction in mannitol excretion following an oral fructose dose. Presently, it is thought that a portion of the orally ingested fructose is converted to a mixture of mannitol and sorbitol, with recycling of the sorbitol back to fructose via the sorbitol dehydrogenase enzyme (K_m for sorbitol is 0.5 mM and for mannitol 50–100 mM), followed by excretion of the mannitol into the urine. It is postulated that the major site of mannitol formation may be the liver.

SUMMARY

This brief review of the sorbitol pathway has attempted to present our current knowledge of the possible role of this accessory pathway of glucose metabolism in the development of some diabetic complications. Clearly hyperglycemia in the diabetic patient is an important factor controlling the activity of the sorbitol pathway. Hyperglycemia in both diabetic patients and experimental animals results in significant accumulations of the products of this pathway in some tissues, and these accumulations have been demonstrated to play an important role in some of these diabetic manifestations. The development of inhibitors of the aldose reductase enzyme affords new means for preventing and treating some of these complications. Nevertheless, we are still hampered by the lack of knowledge of the normal role of this pathway in tissue metabolism. Many technical problems still exist concerning sensitive and specific assays for the products of the sorbitol pathway in tissue studies, as well as with valid techniques for the measurement of the activity of this pathway in clinical situations. It is hoped that clinical studies with aldose reductase inhibotors in the future will further clarify the importance of this accessory pathway of glucose metabolism in diabetes.

Literature Cited

1. United States Department of Health, Education, and Welfare, Public Health Service, Division of Chronic Diseases. 1964. *Diabetes Source Book.* Washington, DC: Public Health Service (Publ. No. 1168)
2. Pirart, J. 1965. *Diabetes* 14:1–9
3. Knowles, H. C., Guest, G. M., Lampe, J., Kissler, M., Skillman, T. G. 1965. *Diabetes* 14:239–73
4. Marble, A. 1967. *Diabetes* 16:825
5. Gregersen, G. 1968. *Diabetologia* 4: 273–77
6. Albisser, A. M. et al 1974. *Diabetes* 23:397
7. Orr, A. P. 1924. *Biochem. J.* 18:171–72
8. Hubbard, R. S., Russell, N. M. 1937. *J. Biol. Chem.* 119:647–61
9. Mann, T. 1946. *Biochem. J.* 40:481–91
10. Mann, T., Parsons, U. 1950. *Biochem. J.* 46:440–50
11. Mann, T. 1964. *Biochemistry of Semen and of the Male Reproductive Tract.* New York: Wiley
12. Hers, H. G. 1956. *Biochim. Biophys. Acta* 22:202–3
13. Hers, H. G. 1960. *Biochim. Biophys. Acta* 37:137
14. Hugget, A. St. G., Warren, F. L., Warren, N. V. 1951. *J. Physiol.* 113:258
15. Hayman, S., Kinoshita, J. H. 1965. *J. Biol. Chem.* 240:877
16. Moonsammy, G. I., Stewart, M. A. 1967. *J. Neurochem.* 14:118–93
17. Gabbay, K. H., Tze, W. J. 1972. *Proc. Nat. Acad. Sci. USA* 69:1435–39
18. Gabbay, K. H. 1972. *Isr. J. Med. Sci.* 8:1626–29
19. Gabbay, K. H., Cathcart, E. S. 1974. *Diabetes* 23:460–68
20. Clements, R. S. Jr., Weaver, J. P., Winegrad, A. I. 1969. *Biochem. Biophys. Res. Commun.* 37:347
21. Clements, R. S., Winegrad, A. I. 1972. *Biochem. Biophys. Res. Commun.* 47:1473–80
22. Kinoshita, J. H., Futterman, S., Satoh, K., Merola, L. O. 1963. *Biochim. Biophys. Acta* 74:340–50
23. Gabby, K. H. 1969. *Diabetes* (Suppl.1) 18:336–37
24. Jedziniak, J. A., Kinoshita, J. H. 1971. *Invest. Ophthalmol.* 10:357–66
25. Mano, Y., Suzuki, K., Yamata, K., Schimazono, N. 1961. *J. Biochem.* 49:618
26. Beutler, E., Guinto, E. 1974. *J. Clin. Invest.* 53:1258–64
27. Blakely, R. L. 1951. *Biochem. J.* 49:257–71
28. McCorkindale, J., Edson, N. L. 1954. *Biochem. J.* 57:518
29. Smith, M. G. 1962. *Biochem. J.* 83:135–44
30. Jedziniak, J. A., Yates, E. M., Kinoshita, J. H. 1973. *Exp. Eye Res.* 99–104
31. Gabbay, K. H. Unpublished
32. Mills, N. C., Means, A. R. 1972. *Endocrinology* 91:147–56
33. Gabbay, K. H. 1974. *Clin. Res.* 22: 468A
34. Kinoshita, J. H. 1965. *Invest. Ophthalmol.* 4:786–99
35. Gabbay, K. H., O'Sullivan, J. B. 1968. *Diabetes* 17:239–43
36. Gabbay, K. H., O'Sullivan, J. B. 1968. *Diabetes* 17:300
37. Kinoshita, J. H., Dvornik, D., Kraml, M., Gabbay, K. H. 1968. *Biochim. Biophys. Acta* 158:472–75
38. Dvornik, D. et al 1973. *Science* 182: 1146–48
39. LeFevre, P. G., Davies, R. I. 1951. *J. Gen. Physiol.* 34:515
40. Wick, A. N., Drury, D. R. 1951. *Am. J. Physiol.* 166:421–23
41. DaCosta, J. M. 1890. *Medical Diagnosis.* Philadelphia: Lippincott. 7th ed.
42. Duke-Elder, W. S. 1925. *Brit. J. Ophthalmol.* 9:167–87
43. Van Heyningen, R. 1959. *Nature (London)* 184:194–95
44. Kinoshita, J. H., Merola, L. O., Satoh, K., Dikmak, E. 1962. *Nature (London)* 194:1085
45. Kinoshita, J. H., Merola, L. O., Dikmak, E. 1962. *Exp. Eye Res.* 1:405
46. Kinoshita, J. H., Merola, L. O. and Dikmak, E. 1962. *Biochim. Biophys. Acta* 62:176
47. Hayman, S., Lou, M. F., Merola, L. O., Kinoshita, J. H. 1966. *Biochim. Biophys. Acta* 128:474
48. Lou, M. F., Kinoshita, J. H. 1967. *Biochim. Biophys. Acta* 141:547
49. Chylack, L. T. Jr., Kinoshita, J. H. 1969. *Invest. Ophthalmol.* 8:401–12
50. Gabbay, K. H., Kinoshita, J. H. 1972. *Isr. J. Med. Sci.* 8:1557–61
51. Stewart, M. A., Passoneau, J. V. 1964. *Biochem. Biophys. Res. Commun.* 17:536
52. Gabbay, H. K., Merola, L. O., Field, R. A. 1966. *Science* 151:209–10
53. Stewart, M. A., Sherman, W. R., Anthony, S. 1966. *Biochem. Biophys. Res. Commun.* 22:488

54. Stewart, M. A. et al 1967. *J. Neurochem.* 14:1057–66
55. Ward, J. D. 1972. *Diabetes* 21:1173–78
56. Stewart, M. A., Passonneau, J. V., Lowry, O. H. 1965. *J. Neurochem.* 12:719–27
57. Abercrombie, M., Johnson, M. L. 1946. *J. Anat.* 80:42
58. Thomas, P. K., Lascelles, R. G. 1965. *Lancet* 1:1355–57
59. Ward, J. D., Barnes, C. G., Fisher, D. J., Jessop, J. D., Baker, R. W. R. 1971. *Lancet* I:428
60. Gabbay, K. H. 1972. *Vascular and Neurological Changes in Early Diabetes,* ed. R. A. Camerini-Davalos, H. S. Cole, 417–24. New York: Academic
61. Gabbay, K. H., Snider, J. J. 1972. *Diabetes* 21:295–300
62. Gabbay, K. H. 1974. *Sugars in Nutrition* (monograph), ed. H. L. Sipple, K. W. McNutt, 470–80. New York: Academic
63. Greene, D. A., deJesus, P. V., Winegrad, A. I. 1974. *Diabetes* (Suppl.1) 23:357
64. Clements, R. S. Jr., deJesus, P. V. Jr., Winegrad, A. I. 1973. *Lancet* I:1137–41
65. deJesus, P. V. Jr. et al 1974. *J. Neurol. Sci.* 21:237
66. Wray, H. L., Winegrad, A. I. 1966. *Diabetologia* 2:82–85
67. Pitkanen, E., Servo, C. 1973. *Clin. Chim. Acta* 44:437–44
68. Spinowitz, B. S., Altszuler, N. A. 1973. *Diabetes* 22:820–24
69. Clements, R. S. Jr., Prockop, L. D., Winegrad, A. I. 1968. *Lancet* 2:384–86
70. Clements, R. S. Jr. et al 1971. *Lancet* 2:671–75
71. Prockop, L. D. 1971. *Arch. Neurol.* 25:126–40
72. Arieff, A. I., Carroll, H. J. 1974. *Diabetes* 23:525–31
73. Assal, J. P., Aoki, T. T., Manzano, F. M., Kozak, G. P. 1974. *Diabetes* 23:405–11
74. Schofield, P. J., Hutton, J. C., Williams, J. F. 1974. *IRCS* 2:1465
75. Morrison, A. D., Winegrad, A. I., Fink, C. J., Lacy, P. E. 1970. *Biochem. Biophys. Res. Commun.* 38:3, 491
76. Corkey, B., Mayhew, D. Unpublished
77. Malaisse, W. J., Sener, A. 1974. *Diabetes* (Suppl.1) 23:378
78. Morrison, A. D., Clements, R. S. Jr., Travis, S. B., Oski, F., Winegrad, A. I. 1970. *Biochem. Biophys. Res. Commun.* 40:1, 199
79. Travis, S. F., Morrison, A. D., Clements, R. S. Jr., Winegrad, A. I., Oski, F. A. 1971. *J. Clin. Invest.* 50:2104–12
80. Clements, R. S. Jr., Morrison, A. D., Winegrad, A. I. 1969. *Science* 166:1007
81. Morrison, A. D., Clements, R. S., Winegrad, A. I. 1972. *J. Clin. Invest.* 51:3114–23
82. Kuck, J. F. R. Jr., Crosswell, H. H. Jr. 1974. *Ophthalmol. Res.* 6:189–96
83. Aloia, J. F. 1973. *J. Lab. Clin. Med.* 82:5
84. Pitkanen, E. 1972. *Clin. Chim. Acta* 38:221–30
85. Blau, K. 1972. *Clin. Chim. Acta* 38:441–47

Copyright 1975. All rights reserved

ANALGESIC ABUSE AND RENAL DISEASE

◆7171

Thomas Murray, M.D. and Martin Goldberg, M.D.
Renal-Electrolyte Section, Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

In 1953 Spühler & Zollinger (1) noted that several patients with chronic interstitial nephritis also ingested large amounts of analgesic compounds. Since that time many hundreds of cases of nephropathy associated with abuse of analgesic drugs have been reported (2). Although experimental approaches have thus far failed to identify the exact pathogenesis, most students of this problem agree that a causal relationship exists between the abuse of analgesics and a characteristic renal disease.

Despite acceptance of this association by most physicians, the number of well documented reports of patients with this form of nephropathy has been relatively low in the United States compared to several other regions of the world, including our North American neighbor, Canada. There are several possible reasons for this apparent discrepancy in regional prevalence, but the most important explanation may be that a substantial number of cases are escaping detection in the United States. If this is true, it should be a cause of major concern, because analgesic abuse nephropathy is preventable. Furthermore, once it develops, cessation of analgesic abuse is often associated with stabilazation of renal function and possibly even improvement.

Although a number of excellent reviews of this subject have been written (2–5), the present one will emphasize the current magnitude of the problem (especially in the USA), review some recently developed concepts on pathogenesis, and outline the information needed to allow physicians to diagnose correctly and treat the cases they encounter.

WHAT IS ANALGESIC ABUSE?

Which Analgesics?

Although a great number of drugs are used for analgesia, those that are commonly considered as possible causes of renal disease in North America (correctly or incorrectly) when abused are aspirin, phenacetin, acetaminophen, caffeine, and codeine. Although the latter two drugs are commonly ingested in concert with the others listed, there is little to support the contention that they contribute significantly to the development of the renal disease (3).

Early reports assumed that phenacetin was the cause of the renal disease because it was the single drug common to the many analgesic mixtures ingested by almost all the well documented cases (1). It was quickly pointed out, however, that such an assumption was unwarranted since most of the reported cases had all ingested the phenacetin in combination with other drugs (3, 5, 6). However, at least one patient has been reported who developed papillary necrosis, a major manifestation of this disease, after the excessive consumption of phenacetin alone (7). A major reason for the difficulty in attributing analgesic abuse nephropathy specifically to phenacetin is that this compound is generally not utilized as a single drug, but more commonly is available in combination with other analgesics, particularly aspirin. Although direct proof that phenacetin alone can cause renal damage in the human is not available, such a conclusion is not ruled out by the clinical data.

Since aspirin is a component of the analgesic mixtures in most of the reported cases, it has been also considered a possible cause of the renal disease. At least five patients have developed papillary necrosis associated with the ingestion of aspirin alone (3, 8). There is, however, a recent study that suggests that the incidence of papillary necrosis and/or renal functional impairment are both extremely rare consequences of the use of aspirin alone, even in very high doses for prolonged periods of time.

It is clear however that the ingestion of excessive amounts of aspirin and phenacetin together, with or without other drugs, has most commonly been reported to be associated with the development of papillary necrosis or renal functional impairment in the absence of other known causes of renal disease. This fact has several implications. It is theoretically possible, for example, that aspirin is the toxic drug but that only those patients who ingest combination analgesics ingest enough aspirin to cause damage. Surveys of patients who have ingested large amounts of aspirin alone, however, make this an unlikely explanation (9). It is also possible that ingestion of one drug modifies the metabolism or alters renal tissue response of the other drug in such manner as to increase its toxicity. Thus, the abuse of the combination of drugs may be necessary to produce the nephropathy. Finally, it is still possible that phenacetin is the primary cause of the renal damage but that its major availability to patients is in combination form.

Acetaminophen (N-acetyl-*p*-aminophenol, APAP, Tylenol®) is presently being used in increasing amounts either alone or in combination with aspirin. Since it is also the major metabolic product of phenacetin in the human (more than 90% of ingested phenacetin is converted to this form within 1 hr), the risk of its excessive use being associated with the development of renal disease should be considered. There are so far relatively few patients reported who have ingested massive amounts of acetaminophen, but there have been three patients described with papillary necrosis after its long term ingestion—two after acetaminophen and aspirin in combination (9, 10), and one after acetaminophen alone (11). Thus, until further experience accumulates to the contrary, this drug, too, should be considered potentially related to the development of kidney disease, at least in combination with aspirin.

What Constitutes Abuse?

The use of aspirin and phenacetin, the latter particularly in combination, is common the world over. Table 1 shows the average per capita consumption of phenacetin in grams per year in various countries as estimated by Gault (12) from production figures. Gross abuse of such drugs, e.g. the intake of 20 or more tablets daily for several years, is easy to define, especially when associated with renal disease or other more obvious signs of drug toxicity. Of course, ingestion of any quantity of drug with dubious clinical indications constitutes abuse also. The question most relevant to this discussion, however, is the minimum amount of analgesic drug consumption (i.e. abuse) that leads to renal damage in a significant number of patients. The precise answer to this question is unknown. Nevertheless, based on studies of large numbers of patients with analgesic-associated renal disease, investigators have attempted to estimate the minimal level of drug abuse that is most likely to be associated with renal disease (2, 13, 14). Definitions of such abuse have been proposed based on the amount consumed per day (e.g. 1 g of phenacetin daily for 1–3 yr) (13) and on the total amount consumed over any period of time (e.g. 3 kg of aspirin or phenacetin) (2, 14). In support of this latter estimate of abuse in relation to renal disease, a Canadian cooperative study suggested that only 1.5% of patients diagnosed as having analgesic abuse nephropathy had ingested less than 1 kg of phenacetin (15). Most of the reported patients with analgesic abuse nephropathy and azotemia, however, have ingested well over the minimum amounts suggested as defining abuse. Various studies have reported the average level of aspirin or phenacetin consumption in these patients to be between 6 and 17 kg (8, 12, 15, 16). Table 2 lists the number of tablets of some of the analgesic compounds commonly available in the United States that a patient would have to ingest to fulfill the minimum abuse criteria.

The frequency of abuse according to any of the proposed criteria has not been well studied in an unselected population. Gault, Rudwal & Redmond found that 6% of male Canadian veterans abused analgesics (defined as an intake of 1 kg of either aspirin or phenacetin) and assumed them to be at risk of developing renal disease

Table 1 Per capita consumption of phenacetin (12)

Country	g/yr
Australia	40
Denmark	25
Switzerland	22
South Africa	12
Scotland	12
U.S.A.	10
Canada	6–7

(17). Eighty-five percent of the patients in this study abused phenacetin and aspirin in combination, while only 15% abused aspirin without phenacetin. Since it has been the general experience that many more females than males abuse analgesics (see Table 3) (2, 8, 13, 18), this estimate of 6% is probably lower than the incidence of abuse in the general Canadian population. Comparable studies of unselected populations in other countries where analgesic drugs are readily available have not been reported. Since, however, the per capita ingestion of analgesics is higher in most of those countries than it is in Canada (see Table 1), it is likely that the prevalence of abuse is also higher.

Table 2 Amount of various analgesic drugs that constitute abuse (expressed in amounts of phenacetin or acetaminophen[a])

Proprietary name	Index drug in combination[a] (Phenacetin)	(Acetaminophen)	No. of tablets to provide 1.0 g/day	No. of years to provide 3.0 kg at 10 tablets/day
Darvon Compound®	+		6	5
Empirin®	+		7	6
Fiorinal®	+		8	7
Norgesic®	+		6	5
SK-65®	+		6	5
Excedrin®		+	11	9
Vanquish®		+	6	5

[a] Phenacetin and acetaminophen are used only as indices of amount of consumption, not necessarily to imply that they are the primary nephrotoxin.

Table 3 Clinical features in patients with analgesic abuse nephropathy

Feature	Survey of Literature (2, 8, 10, 13, 18, 19)	20 cases from Hospital of U. of Pa. (26)
Age	29–72 yr	32–62
Females	50–85%	85%
Headache	35–100%	80%
Anemia	60–90%	65%
Gastrointestinal symptoms	40–60%	40%
Hypertension	15–70%	50%
Urinary tract infection	30–60%	15%
Abnormal urogram	85–95%	95%
Papillary necrosis	25–85%	25%
Pyuria	50–100%	95%

PROOF OF ASSOCIATION OF ABUSE WITH RENAL DISEASE

Proof that the ingestion of analgesics can result in chronic renal disease should ideally include both epidemiological and direct experimental evidence. So far, conclusive evidence for the specific etiological agent(s) is lacking. Therefore, precise knowledge of the toxicological-biochemical mechanisms is not available. On the other hand, several pertinent experimental observations have recently been made which may provide insights into the pathological mechanisms involved. A brief commentary on the problems with the experimental data and a selected review of such data will be presented in the section on pathogenesis.

Clinical and epidemiological attempts to prove a causal relationship have thus far centered around three questions 1. Is there an increased incidence of renal disease in people who abuse analgesics? 2. Is the incidence of analgesic abuse higher in patients with renal disease or certain forms of renal disease than it is in the general population? 3. Is there a correlation between the magnitude of analgesic abuse and (*a*) the frequency of renal disease in a given population or (*b*) the severity of the renal disease in individual patients within that population?

The exact frequency with which renal disease develops in analgesic abusers is difficult to state because the definition of abuse and the criteria used to document renal disease have differed in various studies. Nordenfelt & Ringertz (19) found that 29% of abusers (>1 kg phenacetin) had evidence of "kidney damage." Larson & Mullen (20) found that 33% of abusers (>1 kg phenacetin) had reduced renal function (creatinine clearance) while only 7% of nonusers of analgesics and 11% of nonabusing users had reduced renal function. Lindenberg and co-workers (21) reported that 90% of 39 patients with a history of "marked" abuse had renal functional impairment, and Sorensen (22), who reviewed the European literature, found an average incidence of nephropathy of 22% among abusers. The lowest incidence of renal disease in abusers was reported by Gault (17). In his study only 3% of abusers had severe renal insufficiency and only 3 of 20 abusers studied in detail had evidence of early renal functional impairment. In all of these studies, however, the incidence of renal disease was greater in abusers than in the general population.

The incidence of analgesic abuse in patients with renal disease also supports a causal association between abuse and the development of such disease. Available data suggest that between 3 and 17% of patients with otherwise unexplained end stage renal disease have abused analgesic drugs (15, 23, 24). In a series of 49 patients from Iceland, 38% with nonobstructive pyelonephritis who were without diabetes mellitus or stones were found to be analgesic abusers (13). We recently reported that 20% of patients with clinically diagnosed chronic interstitial nephritis at the Hospital of the University of Pennsylvania abused analgesics, and this was considered to be the most likely primary cause of their disease (see Table 3) (25). All of these studies suggest that patients with renal disease, especially those with the functional and morphological characteristics of chronic tubular-interstitial renal disease (interstitial nephritis, "pyelonephritis"), are much more commonly abusers of analgesics then are persons without renal disease.

Finally, correlations between the incidence or severity of renal disease and the amount of analgesic abused or consumed have also been suggested. A survey by Nordenfelt (26) from Sweden led him to suggest that a decrease in the availability of phenacetin-containing analgesic compounds for a six-year period was associated with a decrease in deaths from renal disease. Another study has provided data suggesting that increasing amounts of analgesic ingestion were associated with increasingly severe degrees of renal functional impairment in patients with analgesic abuse nephropathy (13). Furthermore, numerous studies have shown that cessation of analgesic use frequently results in stabilization of the degree of functional impairment, while continued abuse or the reinstitution of abuse is usually associated with progressive deterioration of renal function (2, 8, 27) (Figure 1).

Despite such extensive clinical data, at least two studies claim to have demonstrated the lack of a correlation between analgesic abuse and renal disease. Sorensen (28) concluded that the relationship was coincidental rather than causal, yet his data showed that the incidence of radiologically demonstrated papillary necrosis was 11% in analgesic users but only 2.3% in nonusers. Lawson (29) found no evidence of an association between a history of analgesic use (as gathered from chart reviews and patient-nurse interviews) and evidence of renal disease in patients admitted to

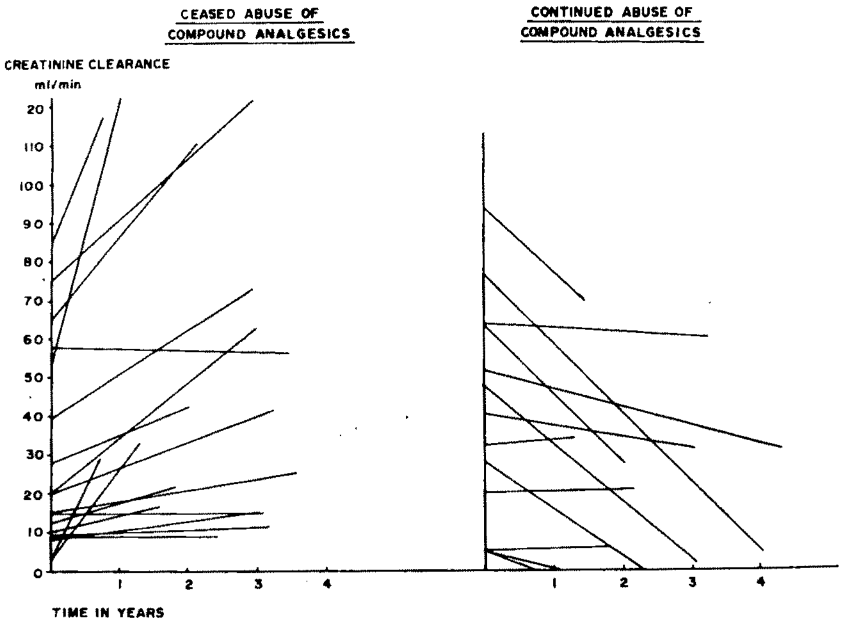


Figure 1 Course of changes in renal function (creatinine clearance) in two groups of patients with analgesic abuse nephropathy: (*a*) group who ceased abuse of analgesic mixtures (*left*) and (*b*) group who continued to abuse analgesic mixtures (*right*). Reprinted with permission from Linton (27).

several hospitals in Boston. This latter study, however, does not provide information critical to answering the causal question, since it is apparent that very few of the patients were abusers of analgesics (by any definition of abuse), and that very few of the analgesic users ingested drugs other than aspirin or acetaminophen alone. As already discussed, the overwhelming majority of reported cases of analgesic abuse nephropathy have involved the ingestion of drug combinations.

TYPE OF RENAL DISEASE

It is clear that the renal disease that occurs in association with analgesic abuse can be classified as interstitial nephritis (see below), rather than glomerulonephritis. These patients, therefore, present to the physician not with evidence of a primary glomerular disorder (in which heavy proteinuria, red cell casts, and nephrotic syndrome are frequent), but instead with symptoms related acutely to obstruction or infection in the urinary tract (often due to papillary necrosis) or more insidiously to those associated with renal insufficiency and/or renal tubular dysfunction.

Debate about the pathology of analgesic abuse nephropathy centers around the nature of the earliest renal lesion and the relationship of this early lesion to the full development of renal disease. The pathologic changes seen in patients with various degrees of renal functional impairment have been summarized by Gault, Blennerhassett & Muehrcke (2). At one extreme, patients with severe functional impairment were found to have severe interstitial fibrosis and periglomerular fibrosis in addition to papillary necrosis. Although varying degrees of glomerular sclerosis were also found in these patients, it was clear that the primary disorder was in the renal interstitium, especially in the renal medulla. At the other extreme, patients with very mild reduction in glomerular filtration rate who were studied by renal biopsy demonstrated morphological changes mainly in the inner medullary area of the kidney (the papillary area was not available for examination). The changes included increased amount of medullary fibrous tissue, focal tubular atrophy, and focal areas of tubular basement membrane thickening. The cortical areas in these patients were essentially normal and there was little cellular infiltration anywhere (in contrast to infectious pyelonephritis). As the degree of functional impairment increased, there was worsening of the medullary changes and the appearance of cortical interstitial nephritis with periglomerular fibrosis. Finally the changes described in those with severe functional impairment were noted. In all the patients in whom the papillae were examined pathologically, there was papillary necrosis (all of these had creatinine clearance values less than 30 ml/min).

Burry (30) and Kincaid-Smith (31) suggest that papillary necrosis is the initial lesion caused by analgesics and that the outer medullary and cortical changes occur only secondarily. Based on their observations, Gault et al (2) on the other hand suggest that medullary changes might occur primarily, with papillary necrosis resulting secondarily, perhaps from interference with the papilla's blood supply. For practical purposes the important point is that the initial pathologic process in analgesic abuse nephropathy begins in the inner medullary region of the kidney (which includes the papilla) and then advances to involve the whole kidney. Furthermore, there is morphological evidence to indicate that the initial changes involve the medullary interstitium, interstitial cells, and tubules rather than the medullary blood vessels, which are involved only secondarily. The final pathologic picture is not unlike that seen with any case of interstitial nephritis (e.g. chronic pyelonephritis) except that papillary necrosis is an almost universal finding.

CLINICAL PICTURE

The patient with analgesic abuse nephropathy may be seen initially by the physician for several different reasons: for complaints concerning the original cause of the abuse (e.g. headache, backache, etc); symptoms or signs as a consequence of progressive renal failure (e.g. anemia, fatigue, uremia, hypertension, etc); acute urinary tract symptoms (e.g. colic, infection, frequency, dysuria, hematuria, etc); manifestations typical of chronic tubular-interstitial renal disease (e.g. nocturia, metabolic acidosis, salt wasting, etc); or an unrelated complaint. Table 3 lists the clinical features that are often noted in these patients and the frequency of their occurrence derived from several large series (2, 6, 8, 10, 18, 25). Most patients are female and most are 40 years of age or older. The stated reason for the abuse is usually headache (for which no objective cause is usually discovered). Other common causes include chronic back pain, arthritis, and use as a general "pick-me-up" (or stimulant), rather than as an analgesic. Many patients have undergone psychiatric treatment, most commonly with a diagnosis of depression. Many also have a history of gastrointestinal disturbances manifested either as gastrointestinal bleeding, ulcers, or chronic dyspepsia, and a significant number have had gastrectomies. Most patients are at least mildly anemic. The anemia appears to be due to a combination of the direct toxic effects of phenacetin and/or acetaminophen on the red blood cell associated with methemoglobinemia, the effect of chronic blood loss, and the effect of uremia on red cell production and destruction (3). Hypertension, which is found in approximately one half of patients, is usually relatively mild and discovered typically after the initial diagnosis of the renal disease and following a long history of analgesic abuse.

In many patients there are no symptoms related to the urinary tract, but a substantial number have a history of recurrent dysuria and/or of past urinary tract infections. A history of renal colic or of the passage of tissue or stones may also be elicited. Although patients who abuse analgesics can also have renal stones, many cases have been described where the stone was shown to be in fact a sloughed papilla or part of a papilla (sometimes calcified) that was passed spontaneously.

Laboratory findings include evidence of azotemia and varying degrees of anemia in all patients. Qualitative examination of the urine usually shows minimal proteinuria (trace to 1+), and a 24-hr protein excretion of 1.0 g is extremely rare. Microscopic examination of the urine sediment may show few abnormalities, but frequently varying degrees of hematuria and pyuria are noted. Only an occasional patient suffers from gross hematuria, and this occurs typically during a bout of colic and concomitant passage of tissue (a sloughed papilla). When present, pyuria may frequently be associated with repeatedly negative urine cultures; nonspecific noncellular casts (hyaline, granular, broad, waxy) are commonly found. White blood cell casts may occasionally be seen, especially with active bacterial infection. Red blood cell casts are not found in the urine sediment. Positive bacterial urine cultures have been reported in a variable percentage of patients (in some reports as high as 60%). It is clear in almost all of the series reported, however, that a substantial number of patients develop azotemia after analgesic abuse without any evidence of past or current infection, many showing multiple negative urine cultures.

Nonspecific changes on *intravenous urography* have been reported in as many as 90% of these patients (13). These are similar to those observed in patients with other forms of interstitial nephritis and include calyceal distortion and scarring along with a striking reduction in renal size late in the course of the disease. Changes diagnostic of papillary necrosis have been found in between 20 and 60% of patients in various series. In some patients the calyceal system is not visualized well enough to allow a definitive comment on the presence of papillary necrosis, while in others such changes clearly are not present. It should be recalled that papillary necrosis can be seen pathologically without having been demonstrated radiographically, so the absence of specific radiologic changes does not rule out its presence (32); therefore, the failure to find such changes by X ray does not militate against the hypothesis that the primary alterations in analgesic abuse nephropathy reside in the papilla.

A troublesome and common characteristic of analgesic abusers is their reluctance to admit to the use of such drugs or to report correctly the extent of their abuse (8, 25). Often these patients do not consider analgesics as medications, so that their use may not be discovered in a routine history. Many of these patients deny outright their analgesic abuse or drastically underestimate the extent of their use even after specific questioning. Seven of the 20 analgesic abusers with nephropathy observed by us initially denied their extensive use of analgesics. The correct history in these 7 patients was obtained only after questioning family members and/or friends (25).

Partially as a consequence of the patient's reluctance to estimate correctly the extent of analgesic use, but also because of the lack of appreciation by the physician of the association between renal disease and analgesic abuse, the correct etiologic diagnosis is often not made, at least initially, in these patients. In our series, the referring physician did not make the correct diagnosis in any of the 20 patients with analgesic abuse nephropathy, and the primary internist caring for the patient at our hospital made the correct diagnosis in only 6 instances (30%) (25).

Another point is that these patients sometimes have a positive family history of analgesic abuse and analgesic abuse nephropathy (25, 26, 33), and this abuse may transcend several generations. At least 2 of our 20 cases had other family members who abused analgesics and had chronic interstitial nephritis.

NATURAL HISTORY OF ANALGESIC ABUSE

It has been well documented that the cessation of analgesic abuse can result in prolonged periods of apparent stabilization of renal function in many patients with analgesic nephropathy (2, 8, 27). This seems to be true particularly in those patients in whom analgesic ingestion can be stopped prior to the end stages of their renal disease. Nevertheless, apparent stabilization has been reported in some patients with far advanced renal failure. In fact, there are some reports suggesting that actual improvement in renal function may follow stoppage of analgesic drugs even in patients with end stage renal disease (34) (see Figure 1). Data on such patients, however, need to be interpreted with some caution. It is possible, for example, that significant reversal of renal dysfunction may have been due to treatment of superimposed infection or relief of urinary tract obstruction. It seems clear, on the other hand, that chronic interstitial nephritis associated with analgesic abuse is not typically a rapidly progressive renal disease (if analgesic consumption is stopped), when compared to chronic glomerular diseases at comparable levels of renal failure.

Conversely, persistence in the abuse of analgesics or the reinstution of their abuse is usually associated with progression of the disease (2, 8, 27). The number of patients who are able to stop their ingestion of analgesics when properly educated is apparently quite high; this makes extensive efforts to educate the patient and the family quite worthwhile. Many of these patients can be convinced to take no medication at all for their original complaint while others may respond favorably to treatment with a minor tranquilizer.

INCIDENCE OF ANALGESIC ABUSE NEPHROPATHY IN THE UNITED STATES

Compared to observations from several other countries (35), analgesic abuse nephropathy was originally felt to be a rare entity in the United States. During the last ten years, however, more than 100 cases have been reported and many more have obviously gone unreported (12, 18). In a recent study of patients from the greater Philadelphia area, we found that 20% of the newly diagnosed cases of interstitial nephritis seen between 1969 and 1972 at the Hospital of the University of Pennsylvania were caused by analgesic abuse (25). From our observations, we estimated that at least 7% of patients with advanced renal disease seen in this region might be associated with analgesic abuse. Since there was nothing uniquely selective about our patient population, it is likely that our estimate is indicative of the general prevalence of this disease in the United States. The main reason for the apparently high incidence of analgesic abuse in our series (compared to previous estimates of the incidence in the U.S.A.) was our increased awareness of the condition in our initial diagnostic evaluation of our patients after admission to the hospital. None of the patients was referred to us with a prior diagnosis of analgesic abuse nephropathy.

PATHOGENESIS

We will not review in detail much of the experimental work that has attempted to define the pathogenesis of the disease as there are other excellent reviews of this kind (2–4). We will, however, try to summarize some of the more important relevant experimental findings. Conflicting results have been published concerning almost all of the areas of investigation. All of the animal studies suffer from a number of shortcomings: e.g. species differences exist with regard to both the metabolism of various drugs and the toxic manifestations of certain compounds, the period of drug ingestion in animals has not been as long (either in total years or in proportion to the animals' lifetime) as it often is in humans, and the doses of various drugs employed have often been extremely high in relation to the amounts consumed daily by the typical patient.

Despite these shortcomings a number of observations in these animal studies may have relevance to the pathogenesis of the disease that occurs in humans. Phenacetin or acetaminophen when given alone has produced papillary necrosis in rats only after large doses have been ingested and then only in a low percentage of animals (36–38). Aspirin alone can cause papillary necrosis experimentally after fairly high doses (39). The combination of aspirin and phenacetin (39–41) has resulted most consistently in the experimental development of papillary necrosis either after high doses given for a short period of time or after doses comparable to those taken by humans given over a long period of time. In addition, in one study, 22% of rats given a combination of aspirin and acetaminophen developed papillary necrosis (39). The superimposition of dehydration upon any of the above drug administration regimens to experimental animals increases the frequency with which papillary necrosis has been produced (40, 41). Thus, when given in high doses, aspirin alone and phenacetin alone may cause renal medullary damage and papillary necrosis. On the other hand, there is good evidence to suggest that there may be something uniquely characteristic about the combination of aspirin and phenacetin that might produce renal damage more consistently than any single agent. In fact, controlled chronic studies in rats recently completed in our laboratory (42) support the concept that renal papillary lesions are produced most consistently by the combination of aspirin and phenacetin rather than by either drug given alone in comparable dosage.

The weight of recent evidence suggests that the renal toxic effect is due to a direct action of one or more metabolites of the analgesic agents. Measurements of renal tissue concentrations of the metabolites of aspirin and phenacetin have recently been reported (43, 44). After the ingestion of either acetaminophen and/or phenacetin by dogs, acetaminophen is concentrated in the kidney within 2 hr and a gradient is developed such that the papilla has the highest and the cortex the lowest concentration of this compound (43). Aspirin has also been shown to concentrate in kidney tissue (both cortex and medulla), but there is conflicting evidence about the existence of a medullary gradient (43–45). Dehydration increases the concentration of both aspirin and acetaminophen in the kidney, and overhydration reduces the concentration of these substances (43).

A hypothesis which relates the site of injury by analgesics to their concentration within the kidney and to their known metabolic effects can be proposed. Since acetaminophen, the major metabolic product of phenacetin, concentrates in the kidney in the renal medulla with the highest concentration at the papillary tip, this could provide an explanation for the characteristic distribution of the primary or initial lesions in analgesic nephropathy. One form of possible tissue injury that this compound might produce is oxidative. For example, in the red blood cell that is deficient in glucose-6-phosphate dehydrogenase (required to produce normal amounts of NADPH), the red cell lipid membranes are uniquely susceptible to oxidative breakdown by certain compounds such as phenacetin and acetaminophen.

Normally, the cells of the kidneys have the ability to defend against such damage through NADPH which is generated via the hexose monophosphate (HMP) shunt. Therefore, oxidative injury might occur only when the supply of NADPH was exhausted either by the accumulation of massive amounts of oxidizing drugs or because NADPH generation was inhibited by some other factor. Salicylate is known to be an inhibitor of the HMP shunt pathway in the red cell, and recently it has been shown to possess a similar action in renal tissue (42). Thus, in the presence of renal tissue accumulation of salicylate, phenacetin or acetaminophen could cause oxidative damage at lower doses than would serve to do this if salicylate were not present. Since it is likely that other substances can serve as oxidants against which the kidney must protect itself, it would not be surprising that the apparent abuse of salicylate "alone" in high enough concentration could result in renal damage in some patients. It should be emphasized that this formulation is merely an hypothesis, and although it fits the known data, there is no overt proof that it is correct. Nevertheless, it can explain why aspirin or phenacetin alone in massive doses sometimes can cause papillary necrosis, but also why the accumulation of these agents together is more likely to produce renal damage.

IMPLICATIONS

The evidence appears to be overwhelming that analgesic abuse can cause structural and functional damage to the kidney. It also seems clear that the magnitude of the problem in many parts of the world including the U.S.A. is such that all possible efforts should be instituted to decrease its frequency. One such effort should be directed at identifying all patients who abuse analgesics and urging the cessation of use of such drugs, especially in those who have already developed evidence of renal involvement. A major step will be accomplished when physicians recognize that the problem is common and initiate serious efforts to obtain an accurate history of all such drugs used by all patients. Further efforts need to be directed to the education of physicians and the public to emphasize the risks in abusing analgesics and to encourage patients to consult knowledgeable physicians for any problem that requires continuous use of these medications. There is of course, no evidence that intermittent use of moderate amounts of analgesic drugs presents any significant danger of renal damage.

Another approach to decreasing the frequency of analgesic abuse nephropathy in those patients who need to take large doses of analgesics for valid medical reasons, or in those analgesic abusers who are unwilling to stop their drug ingestion, might be to restrict the use of analgesic combinations containing a mixture of aspirin and phenacetin. There are at least theoretical reasons to believe that this combination may be associated with an increased hazard of renal damage in abusers (see above). Furthermore, currently available information suggests that combination tablets containing phenacetin offer no notable objective advantages over aspirin alone (46–48). Although similar clinical data are not available concerning aspirin and acetaminophen combinations, there is no a priori reason to suppose that this drug combination would offer any remarkable advantage over either drug alone. In fact, the question should be raised whether it is appropriate to continue to allow widespread availability of analgesic drug combinations containing aspirin and phenacetin or acetaminophen (at least without a physician's prescription) unless and until substantial therapeutic benefits from these combinations compared to the efficacy of the individual drugs alone can be demonstrated.

On the other hand, we find little evidence to recommend the restriction of the use of pure aspirin, pure phenacetin, or pure acetaminophen for valid medical reasons. During treatment with any of these analgesics, patients should be encouraged to ingest increased amounts of fluids, since the renal tissue concentration of metabolites of all these drugs is reduced in the hydrated state.

ACKNOWLEDGMENTS

Portions of the work summarized in this paper were supported by grants from the John A. Hartford Foundation, the National Heart and Lung Institute (HL-00340), and by a training grant from the National Institute of Arthritis, Metabolic and Digestive Diseases (AMO5634).

Literature Cited

1. Spuhler, O., Zollinger, H. 1953. *Z. Klin. Med.* 151:1–12
2. Gault, M., Blennerhassett, J., Muehrcke, R. 1971. *Am. J. Med.* 51:740–56
3. Shelley, J. 1967. *Clin. Pharmacol. Ther.* 8:427–71
4. Abel, J. 1971. *Clin. Pharmacol. Ther.* 12:583–98
5. Kincaid-Smith, P. 1967. *Med. J. Aust.* 2:320–21
6. Gilman, A. 1964. *Am. J. Med.* 36:167–73
7. Lindenberg, O. 1958. *Phenacetin Poisoning and Renal Disease, Study in Honorem, N. I. Nissen.* Copenhagen: Univ. Copenhagen. 70 pp.
8. Murray, R., Lawson, D., Linton, A. 1971. *Brit. Med. J.* 1:479–82
9. New Zealand Rheumatism Association Study. 1974. *Brit. Med. J.* 1:17–21
10. Prescott, L. 1966. *Lancet* 2:1143–45
11. Krikler, D. 1967. *Brit Med. J.* 2:615–16
12. Gault, M., Rudwal, T., Engles, W., Dosseter, J. 1968. *Ann. Intern. Med.* 68:906–25
13. Olafsson, O., Gudmundsson, K., Brekkan, A. 1966. *Acta. Med. Scand.* 179:121–28
14. Schreiner, G. 1965. *Progress in Pyelonephritis,* ed. E. Kross. Philadelphia: Davis. 347 pp.
15. Wilson, D. 1972. *Can. Med. Assoc. J.* 107:749–51
16. Danshorn, J., Fairley, K., Kincaid-Smith, P., King, W. 1966. *Quart. J. Med.* 35:69–93
17. Gault, M., Rudwal, T., Redmond, N. 1968. *Can. Med. Assoc. J.* 98:619–26
18. Fellner, S., Tuttle, E. 1969. *Arch. Intern. Med.* 124:379–82
19. Nordenfelt, O., Ringertz, N. 1959. *Acta. Med. Scand.* 165:321–34
20. Larson, K., Mullen, C. 1959. *Acta. Med. Scand.* 164:53–86
21. Lindenberg, O., Fischer, S., Pedersen, J., Nissen, N. 1959. *Acta. Med. Scand.* 165:321–35
22. Sorensen, A. 1965. See Ref. 14, p. 312
23. Parsons, F., Clark, P., Spoek, M. 1970. *Proc. Eur. Dial. Transplant Assoc.* 7:15–22
24. Shiel, A., Stewart, J., Johnson, J. 1969. *Lancet* 2:917–20
25. Murray, T., Goldberg, M. 1974. *Ann. Intern. Med.* In press
26. Nordenfelt, O. 1972. *Acta. Med. Scand.* 191:11–14
27. Linton, A. 1972. *Can. Med. Assoc. J.* 107:746–48
28. Sorensen, A. 1966. *Nephron* 3:366–76
29. Lawson, D. 1973. *J. Chronic Dis.* 26:39–45
30. Burry, A. 1967. *Nephron* 5:185–201
31. Kincaid-Smith, P. 1967. *Lancet* 11:860–62
32. Lalle, A. 1972. *Am. J. Roentgenol. Radium Ther. Nucl. Med.* CXIV:741–45

33. Grimland, K. 1963. *Acta. Med. Scand. Suppl.* 405:1–26
34. Balter, P., Morris, A., Muchrcbe, R. 1974. *Ann. Intern. Med.* In press
35. Schreiner, G. 1962. *Ann. Intern. Med.* 57:1047–52
36. Fordham, C., Huffines, W., Welt, L. 1965. *Ann. Intern. Med.* 62:738–43
37. Claussen, E. 1964. *Lancet* 2:123–24
38. Nanra, R., Kincaid-Smith, P. 1973. *International Symposium of Problems of Analgesic Abuse,* ed H. Haschek, 67–88. Vienna:Facta Publ.
39. Nanra, R., Kincaid-Smith, P. 1970. *Brit. Med. J.* 3:559–61
40. Abraham, E., Rubenstein, A., Levin, N. 1963. *Nature (London)* 200:695–96
41. Saker, B., Kincaid-Smith, P. 1969. *Brit. Med. J.* 1:161–62
42. Goldberg, M., Myers, C. English, P., Peschel, W. 1971. *J. Clin. Invest.* 50:37a
43. Bluemle, L., Goldberg, M. 1965. *J. Clin. Invest.* 47:2507–13
44. Gault, M. 1970. *Proc. Am. Soc. Nephrol.,* p. 28
45. Quintanilla, A., Kessler, R. 1973. *J. Clin. Invest.* 52:3143–53
46. Beaver, W. 1966. *Am. J. Med. Sci.* 251:576–99
47. Murray, W. 1967. *J. Clin. Pharmacol.* 7:150–55
48. Hill, R., Turner, P. 1969: *J. Clin. Pharmacol.* 9:321–27

Copyright 1975. All rights reserved

BIOCHEMICAL ASPECTS OF SCHIZOPHRENIA

❖7172

Steven Matthysse, Ph.D. and Joseph Lipinski, M.D.
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts 02114

The evidence for and against biochemical theories of schizophrenia has been the subject of several comprehensive reviews (1–3). In this article we will selectively discuss a number of lines of investigation which have attracted considerable recent interest. Research on the biochemistry of schizophrenia may, broadly speaking, be divided into two schools of thought: the "inborn errors of metabolism" model and the pharmacological model. The first derives from the strong evidence for genetic factors in the predisposition to schizophrenia (4–7) and the large number of genetic diseases in which a generalized metabolic disturbance or circulating toxic factor can now be identified. The pharmacological model, on the other hand, starts from the premise that understanding the mechanism of action of antipsychotic drugs may be expected to shed light on the pathophysiology of schizophrenia. The underlying assumption that phenothiazines and butyrophenones are not merely sedatives or "tranquilizers," but act on the core symptoms of schizophrenia, is well supported by a number of clinical studies (8–12).

On the whole, the search for metabolic abnormalities or circulating toxins has been disappointing. While many abnormalities have been reported, typically they have not been independently confirmed or appear to be artifacts of drug therapy, diet, stress, inactivity, or chronic infection (1). Several of the more intriguing hypotheses concerning abnormalities detectable in peripheral fluids will be discussed: the autoimmunity theory, "S-protein," trans-3-methyl-2-hexenoic acid, creatine phosphokinase, histamine, and platelet monoamine oxidase.

THE AUTOIMMUNITY THEORY

In 1967, Robert Heath advanced the hypothesis that schizophrenia is an autoimmune disease. The theory rested on three experiments. In the first, fluorescein-tagged anti-human gamma globulin was used to visualize antibodies bound to postmortem brain sections. Fluorescence was observed on neural cell nuclei of septum and ventral part of the head of the caudate nucleus in brain specimens from schizophrenics, whereas no antibody was detected on tissues of nonschizophrenic patients (13). In the second experiment, serum immune globulins of schizophrenic patients, injected intraventricularly in monkeys, caused the animals to become "dazed and out-of-contact," and their EEGs showed abnormalities in the septal region (14). In the third, antibodies prepared against monkey septal region and caudate nucleus produced similar behavioral and EEG changes when introduced intraventricularly in other monkeys (15).

As yet there has been no confirmation of the presence of anti-brain antibodies in brain tissue from schizophrenics. One group was unable to find any precipitin reaction between sera of psychotic patients and homogenates of normal or schizophrenic brain (16); another did not detect anti-brain antibody even though they also used immunofluorescent staining (17); a third observed cytoplasmic rather than nuclear fluorescence which was not specific for schizophrenia (18). The absence of conspicuous tissue reaction in the nervous system in schizophrenics does not seem favorable to an autoimmune origin.

"S-PROTEIN"

Frohman and Gottlieb (19, 20) have proposed that a low density serum lipoprotein (the "S-protein," an alpha-2 globulin) is fundamentally involved in the etiology of schizophrenia. This protein is said to have a tertiary structure different in schizophrenics and normals. When it is in an alpha-helical configuration, it promotes the uptake of tryptophan into cells, where the tryptophan is postulated to be converted into the psychotomimetic substance N,N-dimethyltryptamine. In a careful study Nicol, Seal & Gottesman (21) found, however, that: (*a*) the low density serum lipoprotein fraction does not increase tryptophan uptake into red blood cells; (*b*) hemolysins likely to be present in sera do have this property; and (*c*) enhancement of tryptophan uptake was negatively, rather than positively, correlated with the diagnosis of schizophrenia among psychiatric patients. The S-protein hypothesis must be regarded as dubious on the basis of the presently available evidence.

TRANS-3-METHYL-2-HEXENOIC ACID

The human olfactory system is a remarkably sensitive chemical detector; indeed, one may recall the decisive role that the odor of phenylpyruvic acid had in the discovery of phenylketonuria. Kathleen Smith became impressed that there was an unusual odor in the back wards of mental hospitals, which could not be accounted for by lack of hygiene. She demonstrated that the odor was not merely an idiosyncratic perception; rats could be trained to distinguish between sweat samples from schizophrenic and nonschizophrenic patients (22). The possibility of microbial origin of the odor was debated (23, 24). In 1969 Smith's hypothesis came into greater scientific prominence, because she and her collaborators identified the odorous substance as trans-3-methyl-2-hexenoic acid by mass spectrometry and nuclear magnetic resonance techniques (25). These investigators have recently found, however, that the presence of this substance in sweat is not specific for schizophrenia (26).

CREATINE PHOSPHOKINASE

A careful series of studies by Herbert Meltzer have investigated the activity of serum creatine phosphokinase (CPK) in acute psychotic patients. He has consistently found the activity of this enzyme to be elevated in acute psychosis of both manic and schizophrenic types (27–29). CPK has two isoenzymes; it is the muscle isoenzyme, not the brain variety, that is elevated (27, 28). Cerebrospinal fluid CPK was found to be normal in acute psychotic patients who had elevated serum CPK (30). Exercise, and even the muscular trauma of a single injection of chlorpromazine, can increase serum CPK, but Meltzer's studies appear to have adequately controlled these factors. It is, rather, the lack of specificity of these changes that leads one to question their significance. In the first place, they occur in acute psychoses in general, rather than in schizophrenia alone. Secondly, CPK elevations have also been reported in a variety of other diseases: muscular dystrophy, myocardial infarction, tetanus, stroke, cerebral infarction, head injury, meningitis, and psychomotor epilepsy (31–36).

More recently, Meltzer and his collaborators have reported histological abnormalities in skeletal muscle biopsies from acute schizophrenics, such as "Z-band spreading" (37, 38). They interpret the alterations as secondary ("neurogenic") consequences of an abnormality in nervous system function (39). Although biopsy samples from schizophrenics can be distinguished from normals using these histological criteria, a sufficiently wide variety of other diseases has not been examined to demonstrate that the abnormality is specific for psychosis.

HISTAMINE

Studies of histamine in schizophrenia have centered around two main areas, blood histamine levels and response to intradermal histamine. A number of investigators have reported that schizophrenics have elevated histamine levels in whole blood, plasma, or serum (40–42). More recently, however, histamine levels in whole blood have been reported to vary in schizophrenics, with some showing elevated, some normal, and some decreased levels (43, 44). Others have noted that schizophrenics manifest extraordinary resistance to intradermal injections of histamine, responding with a much smaller wheal than non-schizophrenic patients or normal controls (45–54). There have also been several studies suggesting that schizophrenics may be less likely to have allergies or to show allergic phenomena than normals, and that occasionally pre-existing allergic states may remit during an acute psychosis (53, 55–61). Since histamine is known to be involved in allergic responses, it has been proposed that the latter findings may indicate an abnormality of histamine metabolism in schizophrenics. These reports have been made more interesting by the findings that histamine occurs in the mammalian nervous system (62), including that of man (63). It has recently been suggested that histamine may act as a central neurotransmitter (64).

It is difficult to interpret the clinical observations in the absence of control for antipsychotic drug treatment (only two of the intradermal histamine studies were conducted in the absence of drugs while in none of the studies of blood histamine levels was there control for drug effect). Elevated blood levels of histamine might be found after the use of phenothiazines such as chlorpromazine and thioridazine which, being also antihistaminics, inhibit catabolism of histamine by histamine methyltransferase (HMT) (65). Chlorpromazine, for example, is capable of increasing significantly the histamine content of animal tissue (62). The diminished response to intradermal histamine might be the result of a primary antihistaminic effect of antipsychotic phenothiazines or tachyphylaxis induced by elevated histamine levels after HMT inhibition. In fact Jodrey has reported that patients receiving chlorpromazine demonstrated increased tolerance to injected histamine (66). To determine if there are abnormalities of histamine metabolism in schizophrenia, sensitive and specific methods of assay for blood levels must be used and strict control of medications must be exercised.

PLATELET MONOAMINE OXIDASE

The first biochemical alteration in schizophrenia which has promise as a genetic marker is the reduction of monoamine oxidase (MAO) in platelets, first found by Murphy and Wyatt (67, 68) and confirmed by Meltzer (69). The decrease does not appear to be an effect of drugs or hospitalization. Manic-depressives had MAO values intermediate between chronic schizophrenics and normals, at approximately the same level as patients with acute schizophreniform psychoses. The fact that this reduction was also observed in identical twins of schizophrenics, whether or not they had the illness, indicates that it may be valuable as a genetic marker (68, 70). These intriguing differences prompted Wyatt and his collaborators to measure MAO in postmortem brain specimens from schizophrenic patients. Two forms of MAO in brain were differentiated by inhibitor and substrate specificities. No significant differences between schizophrenics and controls were found in either form (71, 72). The pathogenic significance of the platelet MAO decrease thus remains obscure, but the phenomenon is certainly interesting and worthy of further study.

THE TRANSMETHYLATION THEORY

This productive hypothesis, which has stimulated nearly two decades of research on the biochemistry of schizophrenia, stands halfway between the inborn error of metabolism and pharmacological models. The initial impetus for the theory was pharmacological. There is a striking similarity in molecular structure between mescaline and the catecholamines dopamine and norepinephrine, and also between N,N-dimethyltryptamine and the indoleamine serotonin. The resemblance between psychotomimetics and naturally occurring neurotransmitters led Harley-Mason to propose that transfer of methyl groups to neurotransmitters along normally excluded pathways might cause autointoxicating substances to be formed (73). The postulated process was all the more intriguing because transmethylation is not only a widely occurring biochemical process, but is important specifically in the natural transformations of catecholamines and indoleamines. Originally the hypothesis focused on methylation of catecholamines—dimethoxyphenethylamine (DMPEA) was the supposed pathogenic agent—but a parallel line of investigation developed in connection with the indoleamines.

Shortly after Harley-Mason's theoretical proposal, the postulated compound DMPEA, was reported to be present in the urine of schizophrenics and not in normal controls (74). Subsequently a voluminous and contradictory literature emerged. One group failed to find DMPEA at all (75); one found it too often—both in schizophrenics and normals (76); one claimed it was an artifact which disappeared when vegetable foods were removed from the diet (77); another group agreed that DMPEA was an artifact, but claimed that animal foods had to be omitted in order to eliminate it (78).

More recent investigations have shown, using the unambiguous identification made possible by mass spectrometry, that DMPEA is a constituent of at least some human urine samples, but it is only one of several substances which probably had been confounded in the original studies (79). DMPEA excretion has also been shown to increase after drinking tea (80). A decisive study, where this definitive chemical analysis is matched by equally careful clinical design, remains to be done.

There are a number of findings which suggest a possible association between the methylated indoleamine N,N-dimethyltryptamine (DMT) and schizophrenia. DMT is a very potent hallucinogen (81), and when administered to schizophrenics can cause an intensification of their symptoms (82). The enzyme indoleamine N-methyltransferase is capable of synthesizing DMT from tryptamine, a natural constituent of human brain (83). This enzyme, originally discovered in rabbit lung (84), has now been reported to be present in a variety of mammalian tissues including human lung (85), blood (86), and brain (87, 88).

Wyatt, Saavedra & Axelrod found that schizophrenics and patients with psychotic depression had higher activity of the tryptamine-methylating enzyme in blood platelets than did alcoholics or normal controls, perhaps because of a dialysable inhibitor found in the nonpsychotic group (86). In studies of monozygotic twins discordant for schizophrenia, it was found that the schizophrenic twins had higher platelet tryptamine-methylating enzyme activities than the discordant twins and that the latter had mean enzyme activities which were not different from 22 normal subjects (89). The higher enzyme activity does not appear, therefore, to be genetically determined. On the negative side, it must also be noted that several studies have attempted to measure DMT levels in whole blood, plasma and/or urine without finding significant differences between controls and schizophrenics (90–92), although it is not possible to rule out rapid metabolism of DMT in the body.

Most biological transmethylations utilize S-adenosylmethionine as a methyl donor, this activated intermediate in turn being formed from methionine by the action of methionine adenosyltransferase. N_5-methyltetrahydrofolate has recently been reported to be the methyl donor for the methylation of dopamine to epinine (93), and has been postulated to be active in the methylation of indoleamines via indolethylamine N-methyltransferase (94–96), but its role is controversial (97, 98). Measurements of blood levels of S-adenosylmethionine (99) and methionine adenosyltransferase (100) have not revealed any abnormalities in schizophrenic patients.

The inconclusive results with DMPEA and the methylated indoleamines, and the negative conclusions regarding methyl donors, would have caused interest to wane in the transmethylation theory, were it not for the observation of Pollin, Cardon & Kety that methionine loading causes exacerbation of schizophrenic psychoses (101). Methionine loading is relevant because, as mentioned above, it is the ultimate source of methyl groups for transmethylation reactions. The effect is unusual among observations on schizophrenia in that it has been confirmed by a number of research groups and disconfirmed by none (102–107). Kakimoto concluded that the symptoms resembled a superimposed toxic psychosis more than an exacerbation of schizophrenia (108). In another study, however, only two of the seven reactions were described as toxic, the other five as more like exacerbations of the psychotic process (109). Methionine has many metabolic and pharmacological actions in addition to methyl group donation, which could account for its psychotogenic effect: it reduces seizure threshold (110), and its metabolites include homocysteine and cystathionine, both of which have been found in some instances to be associated with psychosis [in homocystinuria (111–114) and cystathioninuria (115)]. The effects of methionine-loading cannot, therefore, be taken as conclusive evidence for the transmethylation theory.

If a methyl donor made schizophrenics worse, it seemed possible that a methyl acceptor might make them better: hence a trial was carried out of nicotinamide, a substance which is capable of accepting methyl groups to become N^1-methylnicotinamide. Hoffer described successful treatment of 13 out of 17 schizophrenics in 3 to 5 days with daily oral nicotinamide, but several groups have not been able to replicate his claim (116–119). Indeed, in one study nicotinamide had a harmful effect: "increased hostility, aggressiveness and irritability" (120). There is little doubt that the treatment does not work (121). On the other, Baldessarini observed that nicotinamide does not effectively lower rat liver or brain S-adenosylmethionine (122), so its failure as a therapeutic agent may not militate against the transmethylation theory.

In view of the absence of scientific evidence for the efficacy of "megavitamin therapy," it is surprising that the treatment is still widely practiced. Four factors may account for its continued usage. (*a*) Nicotinamide is presumed to be harmless, although in the large doses that are prescribed there may be danger to patients with ulcer, gout, diabetes, or liver disease (123). Pigmented hyperkeratosis has also been reported (124). (*b*) Advocates of nicotinamide therapy have generally avoided double-blind design on the grounds that the vitamin treatment has to be individualized for each patient. (*c*) The treatment plan usually consists of adding large doses of vitamins to the medically recognized standard therapy, but favorable results are often attributed to the vitamins. (*d*) Vitamin therapy was also proposed by Pauling (125) for a different reason, namely, that schizophrenics might be vitamin-deficient in the brain even though dietary intake was normal. The rationale was that loss or partial reduction of the capacity to synthesize vitamins would have a selective advantage in an environment where they were abundant; mutation to restore the synthetic machinery would be rare. At the present time the vitamin deficiency theory is highly conjectural (121).

Although the nicotinamide therapy offshoot of the transmethylation hypothesis has not been successful, the theory continues to be important, especially with respect to the indoleamines, in view of the existence of an enzyme capable of synthesizing the psychotomimetic substance N,N-dimethyltryptamine. Carefully controlled studies of cerebrospinal fluid and postmortem brain specimens, using mass spectrometric techniques for biochemical identification, are needed to resolve the controversy.

THE DOPAMINE HYPOTHESIS

The pharmacological approach to schizophrenia has been dominated, in the past few years, by interest in the hypothesis that phenothiazine and butyrophenone tranquilizers block dopamine receptors in the brain (for reviews, see references 126–128). The evidence is necessarily indirect, since the "dopamine receptor" has not been isolated and characterized as a molecule. Closest to a direct demonstration of receptor blockade is the work of Greengard's laboratory on dopamine-stimulated adenylate cyclase. This enzyme, present in homogenates of rat caudate nucleus, increases in activity in the presence of dopamine. The facilitatory effect is blocked by chlorpromazine and haloperidol (129). Direct neurophysiological evidence of dopamine blockade is difficult to demonstrate, because when phenothiazines are introduced into the vicinity of a neuron they have an anesthetic action. Responses of striatal neurons to iontophoretic dopamine, however, are blocked by local application of chlorpromazine, even after the anesthetic effect wears off (130).

The remaining evidence for blockade of dopamine synapses by phenothiazine and butyrophenone tranquilizers is more indirect than the adenylate cyclase and iontophoretic studies, but taken as a whole it is compelling. The rate of turnover of dopamine in the brain (synthesis and destruction, which are equal in the steady state) is increased by antipsychotic drugs (131–134). Turnover is thought to rise because dopaminergic systems (at least the nigrostriatal tract) increase in activity to compensate for the synaptic blockade. There is evidence for a descending inhibitory pathway from the striatum to the substantia nigra, which could function in this way as a feedback control mechanism (135). Feedback control within the synapse, mediated by presynaptic receptors that inhibit the release of dopamine (136), may operate in the same direction. There is also some interesting neurophysiological evidence in support of the hypothesis. Systemic administration of amphetamine causes a reduction in the firing rate of neurons in the substantia nigra, presumably by stimulating the negative feedback control system; chlorpromazine restores the firing rate to normal (137).

Clinical observations also support the hypothesis of dopamine blockade. Motor side effects reminiscent of the symptoms of Parkinson's disease are a major complication of antipsychotic drugs; and in Parkinson's disease there is a deficiency of brain dopamine. Conversely, L-dopa cannot be used to counteract the extrapyramidal side effects because it makes the psychosis more florid (138). It is interesting that anticholinergic agents can be used to treat the parkinsonian side effects of antipsychotic drugs without making the psychosis worse (although not all agree; see reference 139); this fact is important in the development of the theory (see below).

Another important source of evidence for the dopamine blockade hypothesis is the action of several drugs which although not specific for dopamine, inhibit or facilitate dopaminergic transmission. Reserpine, alpha-methyltyrosine, methylphenidate, and amphetamine are in this category. Reserpine, which interferes with vesicular storage of biogenic amines, was the first effective chemotherapy of schizophrenia, and alpha-methyltyrosine, which inhibits catecholamine synthesis, is reported to decrease substantially the dose of phenothiazines required for treatment (140). Methylphenidate and amphetamine, which facilitate catecholamine transmission, are both psychotomimetic under certain conditions: methylphenidate exacerbates existing psychotic states (141) and amphetamine is capable of eliciting a paranoid psychosis in nonpsychotic individuals (142, 143).

Starting from the blockade of dopamine transmission by antipsychotic drugs, it is possible to formulate hypotheses of the etiology of schizophrenia in a number of ways. The most obvious is that too much dopamine is released at synapses in the central nervous system. Alternatively, dopamine receptors may be hypersensitive and react excessively to a normal quantity of released dopamine. An antagonistic system [perhaps cholinergic (144)] may be underactive. The conversion of dopamine to noradrenaline could be impaired, causing a relative dopaminergic excess and noradrenergic deficiency. There might be a defect in a feedback pathway controlling a branch of the dopamine system. Thus there are several "dopamine hypotheses" of the etiology of schizophrenia corresponding to the premise that antipsychotic drugs block dopamine synapses.

There is a subtle shift of emphasis in passing from the pharmacological to the etiological level, which poses a logical problem, more serious than the fact that there are several alternative formulations. There is not much doubt that dopamine blockade is *an* action of phenothiazines; the question is whether it is *the* action of phenothiazines; that is, whether of the many actions of these drugs dopamine blockade is the one which underlies their antipsychotic properties.

There is some reason to think the logical step can be made. In the first place the dopamine-blocking effects occur in vivo at clinically relevant doses and in vitro at reasonable concentrations. Secondly, nonantipsychotic phenothiazines (closely related chemically, but used as antihistaminics, anticholinergics, or sedatives) generally do not share the dopamine turnover enhancing action of the antipsychotics (128). Thiethylperazine, an antiemetic phenothiazine, is an exception because it does have a marked effect on dopamine turnover, although it is not used in the treatment of schizophrenia. It is possible, however, that its antipsychotic properties have not been thoroughly tested.

Thioridazine is generally agreed to have a lower incidence of parkinsonian side effects (145), and also to have a weaker effect on dopamine turnover (128), than its therapeutic efficacy would predict. It appears that this discrepancy may be accounted for, however, by the interaction of the cholinergic system. As mentioned before, anticholinergic agents can be used to counteract the extrapyramidal side effects of tranquilizing drugs without destroying their antipsychotic action. This is one of several lines of evidence that led to the hypothesis of a cholinergic-dopaminergic balance in the corpus striatum (144, 146). According to recent experiments by Snyder, thioridazine has a much stronger anticholinergic action in the central nervous system than other phenothiazines of similar antipsychotic potency (147). To complete this formulation it is necessary to assume that in those regions of the brain involved in the pathophysiology of schizophrenia, a dopaminergic-cholinergic balance does *not* prevail. Actually there is some evidence that this is the case. The reduction in firing rate of substantia nigra neurons by amphetamine (see above) is antagonized by chlorpromazine but not by thioridazine, whereas when the same experiment is carried out in the region which contains the cells of origin of the dopaminergic tract innervating the limbic system, the two drugs have an equivalent action (148). Anticholinergic agents antagonize the dopamine turnover enhancing effect of haloperidol in the striatum but apparently not in the limbic system (149). Clozapine, which like thioridazine has antipsychotic actions without strong extrapyramidal side effects, causes a greater increase in dopamine turnover in the limbic system than in the caudate (150).

A question which remains to be clarified is whether tolerance develops to the dopamine-blocking effects of tranquilizing drugs. The antipsychotic actions of phenothiazines do not diminish after repeated dosage, whereas tolerance formation is characteristic of many of their other pharmacological actions. One would expect, therefore, that if dopamine blockade accounts for the antipsychotic properties of phenothiazines, it ought not to decrease after chronic administration. Some authors find tolerance (151, 152); others do not (153).

The most direct formulation of the dopamine hypothesis referred to above, excessive release of dopamine at central synapses, is not supported by the clinical studies that have been done. The major metabolite of dopamine, homovanillic acid, is not elevated in the cerebrospinal fluid of schizophrenics (154–156). Prolactin release is under inhibitory control by dopaminergic neurons, so that serum prolactin might be expected to be lower than normal if dopamine release is abnormally high; but no significant difference in schizophrenia was found (157). It is, of course, conceivable that dopaminergic systems other than those of the striatum and hypothalamus, which are primarily reflected by these two measurements, are the ones affected.

The formulation of the dopamine hypothesis in terms of diminished conversion of dopamine to noradrenaline is given some support by studies of Wise, Baden & Stein, who reported a profound decrease in dopamine beta-hydroxylase in brains of schizophrenics as compared with controls (158). This observation could not be confirmed by Wyatt and his collaborators, who concluded that the results (they observed a nonsignificant trend in the expected direction) could be accounted for by differential postmortem degeneration (159). Dopamine beta-hydroxylase is not decreased in the plasma of schizophrenic patients (100, 160).

Even though the experimental evidence for these two formulations of the dopamine hypothesis is not convincing, it is well to recall that there are other possibilities, such as hypersensitivity of dopamine receptors or underactivity of an antagonistic system. It is also conceivable that the regional extent of the neurochemical abnormality is too small for it to be detectable in peripheral fluids or by gross postmortem analysis. The indirect evidence from the blocking action of antipsychotic drugs is not sufficient in itself to establish the etiological significance of dopamine, but as one of the first good clues to the pathophysiology of this baffling disease it is not likely to be abandoned easily.

Literature Cited

1. Kety, S. S. 1967. Current biochemical approaches to schizophrenia. *N. Engl. J. Med.* 276:325–31
2. Weil-Malherbe, H., Szara, S. I. 1971. *The Biochemistry of Functional and Experimental Psychoses.* Springfield, Ill.
3. Wyatt, R. J., Termini, B. A., Davis, J. 1971. Biochemical and sleep studies of schizophrenia: a review of the literature —1960–1970. Part I. Biochemical studies. *Schizophrenia Bull.* 4:10–66
4. Kety, S. S., Rosenthal, D., Wender, P. H., Schulsinger, F. 1968. The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. *J. Psychiat. Res.* 6: 345–62
5. Heston, L. L., Denney, D. 1968. Interactions between early life experience and biological factors in schizophrenia. In *The Transmission of Schizophrenia,* ed. D. Rosenthal, S. S. Kety, 363–76. Oxford: Pergamon
6. Pollin, W. 1972. The pathogenesis of schizophrenia—possible relationships between genetic, biochemical and experiential factors. *Arch. Gen. Psychiat.* 27:29–37
7. Allen, M. G., Cohen, S., Pollin, W. 1972. Schizophrenia in veteran twins: a diagnostic review. *Am. J. Psychiat.* 128:939–45
8. Davis, J. M. 1965. Efficacy of the tranquilizing and antidepressant drugs. *Arch. Gen. Psychiat.* 13:552–72
9. Klerman, G. L., Davidson, E. M., Kayce, M. M. 1964. Factors influencing the clinical responses of schizophrenic patients to phenothiazine drugs and to placebo. *Psychiat. Res. Rep. Am. Psychiat. Assoc.* 19:97–113
10. N.I.M.H. Psychopharmacology Service Center Collaborative Study Group. 1964. Phenothiazine treatment in acute schizophrenia. *Arch. Gen. Psychiat.* 10:246–61
11. Goldberg, S. C., Klerman, G. L., Cole; J. O. 1965. Changes in schizophrenic psychopathology and ward behavior as a function of phenothiazine treatment. *Brit. J. Psychiat.* 111:120–33
12. Klein, D. R., Davis, J. M. 1969. *Diagnosis and Drug Treatment of Psychiatric Disorders.* Baltimore: Williams & Wilkins
13. Heath, R. G., Krupp, I. M. 1967. Schizophrenia as an immunologic disorder. I. Demonstration of antibrain globulins by fluorescent antibody techniques. *Arch. Gen. Psychiat.* 16:1–9
14. Heath, R. G., Krupp, I. M., Byers, L. W., Liljekvist, J. I. 1967. Schizophrenia as an immunologic disorder. II. Effects of serum protein fractions on brain function. *Arch. Gen. Psychiat.* 16:10–23
15. Heath, R. G., Krupp, I. M., Byers, L. W., Liljekvist, J. I. 1967. Schizophrenia as an immunologic disorder. III. Effects of antimonkey and antihuman brain antibody on brain function. *Arch. Gen. Psychiat.* 16:24–33
16. Rubin, R. T. 1965. Investigation of precipitins to human brain in sera of psychotic patients. *Brit. J. Psychiat.* 111:1003–6
17. Wittingham, S., MacKay, I. R., Jones, I. H., Davies, B. 1968. Absence of brain antibodies in patients with schizophrenia. *Brit. Med. J.* 1:347–48
18. Boehme, D. H., Cottrell, J. C., Dohan, C., Hillegass, L. M. 1973. Fluorescent antibody studies of immunoglobulin binding by brain tissues. *Arch. Gen. Psychiat.* 28:202–7
19. Gottlieb, J. S., Frohman, C. E., Beckett, P. G. S. 1969. A theory of neuronal malfunction in schizophrenia. *Am. J. Psychiat.* 126:149–56
20. Harmison, C. R., Frohman, C. E. 1972. Conformational variation in a human plasma lipoprotein. *Biochemistry* 11: 4985–93
21. Nicol, S., Seal, U. S., Gottesman, I. I. 1973. Serum from schizophrenic patients: effect on cellular lactate stimulation and tryptophan uptake. *Arch. Gen. Psychiat.* 29:744–51
22. Smith, K., Sines, J. O. 1960. Demonstration of a peculiar odor in the sweat of schizophrenic patients. *Arch. Gen. Psychiat.* 2:184–88
23. Posner, H. S., Culpan, R., Stewart, A. 1962. Cause of the odor of a schizophrenic patient. *Arch. Gen. Psychiat.* 7: 108–13
24. Skinner, K., Smith, K., Rich, E. 1964.

Bacteria and the "schizophrenic odor." *Am. J. Psychiat.* 121:64–65

25. Smith, K., Thompson, G. F., Koster, H. D. 1969. Sweat in schizophrenic patients: identification of the odorous substance. *Science* 166:398–99
26. Gordon, S. G., Smith, K., Rabinowitz, J. L., Vagelos, P. R. 1973. Studies of trans-3-methyl-2-hexenoic acid in normal and schizophrenic humans. *J. Lipid. Res.* 14:495–503
27. Meltzer, H. 1968. Creatine kinase and aldolase in serum: abnormality common to acute psychoses. *Science* 159:1368–70
28. Meltzer, H. 1969. Muscle enzyme release in the acute psychoses. *Arch. Gen. Psychiat.* 21:102–12
29. Meltzer, H., Elkun, L., Moline, R. A. 1969. Serum-enzyme changes in newly admitted .psychiatric patients. *Arch. Gen. Psychiat.* 21:731–38
30. Martin, W. A., Garey, R. E., Heath, R. G. 1972. Cerebrospinal fluid creatine kinase in acutely psychotic patients. *J. Neurol. Neurosurg. Psychiat.* 35:726–29
31. Cunningham, L. A., Rich, C. L., Woodruff, R. A. Jr., Olney, J. W. 1974. Creatine phosphokinase and psychiatric illness. *Brit. J. Psychiat.* 124:87–91
32. Acheson, J., James, D. C., Hutchinson, E. C., Westhead, R. 1965. Serum-creatine-kinase levels in cerebral vascular disease. *Lancet* 1:1306–7
33. Dubo, H., Park, D. C., Pennington, R. J. T., Kalbag, R. M., Walton, J. N. 1967. Serum-creatine-kinase in cases of stroke, head injury, and meningitis. *Lancet* 2:743–48
34. Belton, N. R., Backus, R. E., Millichap, J. G. 1967. Serum creatine phosphokinase activity in epilepsy. *Neurology* 17:1073–76
35. Brody, I. A., Hatcher, M. A. 1967. Origin of increased serum creatine phosphokinase in tetanus. *Arch. Neurol. (Chicago)* 16:89–93
36. Goto, I., Katsuki, S. 1970. Creatine phosphokinase isoenzymes in pathological human serum. *Clin. Chim. Acta.* 30:795–99
37. Engel, W. K., Meltzer, H. 1970. Histochemical abnormalities of skeletal muscle in patients with acute psychoses. *Science* 168:273–75
38. Meltzer, H., Engel, W. K. 1970. Histochemical abnormalities of skeletal muscle in acutely psychotic patients. *Arch. Gen. Psychiat.* 23:492–502
39. Meltzer, H. Y., Crayton, J. W. 1974. Subterminal motor nerve abnormalities in psychotic patients. *Nature* 249: 373–75
40. Cassell, W. A., Newton, G., Cho, M. 1967. Serum histamine concentrations in schizophrenic and non-schizophrenic patients. *J. Schizophrenia* 7:65–73
41. Lovett-Doust, J. W., Husdin, H., Salna, M. E. 1956. Blood histamine and tissue cell anoxia in mental disease. *Nature* 178:492
42. Stern, P., Hukovic, S., Madjerek, Z., Karabaic, S. 1957. Histamine-gehalt im blut von schizophrenikern (Histamine levels in the blood of schisophrenics). *Arch. Int. Psychodyn.* 109:294–99
43. Pfeiffer, C. C., Ilien, V., Goldstein, L., Jenney, E. H. 1969. *Serum polyamine levels in schizophrenia and other objective criteria of clinical status. In: Schizophrenia, Current Concepts and Research, ed. D. V. S. Sankar, 557–63. Hicksville, N.Y.: PJD Publ.*
44. Pfeiffer, C. C. 1972. Blood histamine, basophile counts and trace elements in the schizophrenias. *Rev. Can. Biol.* 31:73–76
45. Sackler, M. D., Sackler, R. R., Sackler, A. M., Van Ophjuisen, J. H. W. 1951. The technique of histamine biochemotherapy and suggestions for its use in psychiatry. *J. Nerv. Ment. Dis.* 113:40–51
46. Ermala, P., Autio, L. 1951. On intradermal histamine tests in schizophrenia. *Acta Psych. Neurol. Scand. Suppl.* 60:136–44
47. Lucy, J. D. 1954. Histamine tolerance in schizophrenia. *Arch. Neurol. Psychiat.* 71:629–39
48. Mazur, W. P. 1955. Histamine tolerance in a schizophrenic. *N. S. Med. Bull.* 34:209
49. Freedman, D. X., Redlich, F. C., Igersheimer, W. W. 1956. Psychosis and Allergy: Experimental Approach. *Am. J. Psychiat.* 112:873–77
50. Weckowicz, T. E., Hall, R. 1958. Skin histamine reaction in schizophrenic and non-schizophrenic mental patients. *J. Nerv. Ment. Dis.* 126:415–20
51. Sackler, A. M., Sackler, R. R., Martí-Ibánez, F., Sackler, M. D. 1954. Contemporary psychodynamic therapeutic trends in psychiatry. *J. Clin. Exp. Psychopathol.* 15:382–400
52. LeBlanc, J., Lemieux, L. 1961. Histamine and mental disease. *Med. Exp.* 4:214–22
53. Simpson, G. M., Kline, N. S. 1961. Histamine wheal formation and mental illness. *J. Nerv. Ment. Dis.* 133:19–24

54. Cowan, M. A. 1969. Behavioral neuropsychiatry. The histamine wheal in normal and psychiatric patients. *Behav. Neuropsychiat.* 1:15–20
55. Funkenstein, D. H. 1950. Psychophysiologic relation of asthma and uriticaria to mental illness. *Psychosom. Med.* 12:377–438
56. Ehrentheil, O. F. 1957. Common medical disorders rarely found in psychotic patients. *Arch. Neurol. Psychiat.* 77: 178–86
57. McAllister, R. M., Hecker, A. D. 1949. The incidence of allergy in psychotic reactions. *Am. J. Psychiat.* 105:843–48
58. Leigh, D., Lovett-Doust, J. W. 1953. Asthma and psychosis. *J. Ment. Sci.* 99:489–96
59. Mandell, A. J., Younger, C. B. 1962. Asthma alternating with psychiatric symptomatology. *Calif. Med.* 96: 251–53
60. Weblin, J. F. 1963. Psychogenesis in asthma: an appraisal with a view to family research. *Brit. J. Med. Psychol.* 36:211–25
61. Cassell, W. A., Fisher, S. 1963. Body-image boundaries and histamine flare reaction. *Psychosom. Med.* 25:344–50
62. Green, J. P. 1970. Histamine. In *Handbook of Neurochemistry,* ed. A. Lajtha, 4:221–50. New York: Plenum
63. Lipinski, J. F., Schaumburg, H. H., Baldessarini, R. J. 1973. Regional distribution of histamine in the human brain. *Brain Res.* 52:403–8
64. Snyder, S. H., Taylor, K. M. 1972. Histamine in the brain: a neurotransmitter? In *Perspectives in Neuropharmacology: A Tribute to Julius Axelrod,* ed. S. H. Snyder, 43–74. New York: Oxford Univ. Press
65. Taylor, K. M., Snyder, S. H. 1972. Histamine methyltransferase inhibition and potentiation by antihistamines. *Molec. Pharmacol.* 8:300–10
66. Jodrey, L. H., Smith, J. A. 1959. Releasable histamine levels and histamine tolerance in the tissues of 291 psychotic patients. *Am. J. Psychiat.* 115:801–7
67. Murphy, D. L., Wyatt, R. J. 1972. Reduced monoamine oxidase activity in blood platelets from schizophrenic patients. *Nature* 238:225–26
68. Murphy, D. L., Belmaker, R., Wyatt, R. J. 1975. Monoamine oxidase in schizophrenia and other behavioral disorders. In *Catecholamines and Their Enzymes in the Neuropathology of Schizophrenia,* ed. S. Matthysse, S. S. Kety, 221–47. Oxford: Pergamon
69. Meltzer, H. Y., Stahl, S. M. 1974. Platelet monoamine oxidase activity and substrate preferences in schizophrenic patients. *Res. Commun. Chem. Pathol. Pharmacol.* 7:419–31
70. Wyatt, R. J. et al 1973. Reduced monoamine oxidase activity in platelets: a possible genetic marker for vulnerability to schizophrenia. *Science* 179: 916–18
71. Schwartz, M. A., Wyatt, R. J., Yang, H.-Y. T., Neff, H. H. 1974. Multiple forms of brain monoamine oxidase in schizophrenic and normal individuals. *Arch. Gen. Psychiat.* 31:557–60
72. Schwartz, M. A., Aikens, A. M., Wyatt, R. J. 1974. Monoamine oxidase activity in brains from schizophrenic and mentally normal individuals. *Psychopharmacologia* 38:319–28
73. Osmond, H., Smythies, J. 1952. Schizophrenia: a new approach. *J. Ment. Sci.* 98:309–15
74. Friedhoff, A. J., Van Winkle, E. 1962. The characteristics of an amine found in the urine of schizophrenic patients. *J. Nerv. Ment. Dis.* 135:550–55
75. Nishimura, T., Gjessing, L. R. 1965. Failure to detect 3,4-dimethoxyphenylethylamine and bufotenine in the urine from a case of periodic catatonia. *Nature* 206:963–64
76. Takesada, M., Kakimoto, Y., Sano, I., Kaneko, Z. 1963. 3,4-dimethoxyphenylethylamine and other amines in the urine of schizophrenic patients. *Nature* 199:203–4
77. Perry, T. L., Hansen, S., MacDougall, L., Schwartz, C. J. 1966. Urinary amines in chronic schizophrenia. *Nature* 212:146–48
78. Von Studnitz, W., Nyman, G. E. 1965. Excretion of 3,4-dimethoxyphenylethylamine in schizophrenia. *Acta Psychiat. Scand.* 41:117–21
79. Creveling, C. R., Daly, J. W. 1967. Identification of 3,4-dimethoxyphenylethylamine from schizophrenic urine by mass spectrometry. *Nature* 216:190–91
80. Stabenau, J. R., Creveling, C. R., Daly, J. 1970. The "pink spot," 3,4-dimethoxyphenylethylamine, common tea, and schizophrenia. *Am. J. Psychiat.* 127: 611–16
81. Szara, A. 1967. Hallucinogenic amines and schizophrenia (with a brief addendum on N-dimethyltryptamine). In *Amines and Schizophrenia,* ed. H. W. Himwich, S. S. Kety, J. R. Smythies, 181–97. Oxford: Pergamon

82. Turner, W. J., Merlis, S. 1959. Effects of some indolealkylamines in man. *Arch. Neurol. Psychiat.* 81:121–29
83. Saavedra, J. M., Axelrod, J. 1973. Effect of drugs on the tryptamine content of rat tissues. *J. Pharmacol. Exp. Ther.* 185:523–29
84. Axelrod, J. 1961. Enzymatic formation of psychotomimetic metabolites from normally occurring metabolites. *Science* 134:343–44
85. Mandel, L., Ahn, H. S., VandenHeuvel, W. J. A., Walker, R. W. 1972. Indoleamine-N-methyltransferase in human lung. *Biochem. Pharmacol.* 21:1197–1200
86. Wyatt, R. J., Saavedra, J. M., Axelrod, J. 1973. A dimethyltryptamine-forming enzyme in human blood. *Am. J. Psychiat.* 130:754–60
87. Mandell, A. J., Morgan, M. 1971. Indole(ethyl)amine-N-methyltransferase in human brain. *Nature New Biol.* 230:85–87
88. Saavedra, J. M., Axelrod, J. 1972. Psychotomimetic N-methylated tryptamines: Formation in brain in vivo and in vitro. *Science* 175:1365–66
89. Wyatt, R. J., Saavedra, J. M., Belmaker, R., Cohen, S., Pollin, W. 1973. The dimethyltryptamine forming enzyme in blood platelets: A study in monozygotic twins discordant for schizophrenia. *Am. J. Psychiat.* 130: 1359–61
90. Wyatt, R. J., Mandel, L. R., Ahn, H. S., Walker, R. W., VandenHeuvel, W. J. A. 1973. Gas-chromatographic-mass spectrometric isotope dilution determination of N,N-dimethyltryptamine concentrations in normal and psychiatric patients. *Psychopharmacologia* 31: 265–70
91. Bidder, T. G., Mandel, L. R., Ahn, H. S., VandenHeuvel, W. J. A., Walker, R. W. 1974. Blood and urinary dimethyltryptamine concentrations in acute psychotic disorders. *Lancet* 1:165
92. Lipinski, J. F., Mandel, L. R., Ahn, H. S., VandenHeuvel, W. J. A., Walker, R. W. 1974. Blood dimethyltryptamine concentrations in psychotic disorders. *Biol. Psychiat.* 9:89–91
93. Laduron, P. 1972. N-methylation of dopamine to epinine in brain tissue using N-methyltetrahydrofolic acid as the methyl donor. *Nature New Biol.* 238:212–13
94. Banerjee, S. P., Snyder, S. H. 1973. Methyltetrahydrofolic acid mediates N- and O-methylation of biogenic amines. *Science* 182:74–75
95. Laduron, P. 1973. N-methylation of biogenic amines in the brain with 5-methyltetrahydrofolic acid as the methyl donor: a possible implication in schizophrenia. In *Advances in Neuropsychopharmacology,* ed. T. A. Ban et al, 235–43. Amsterdam: North-Holland
96. Hsu, L. L., Mandell, A. J. 1973. Multiple N-methyltransferases for aromatic alkylamines in brain. *Life Sci.* 13: 847–58
97. Lin, R. -L., Narasimhachari, N. 1974. Evidence for the absence of amine-N-methylation and O-methylation in indoleamines with methyltetrahydrofolic acid-dependent N-methyltransferase. *Res. Commun. Chem. Pathol. Pharmacol.* 8:535–42
98. Barchas, J. D. Personal communication
99. Matthysse, S., Baldessarini, R. J. 1972. S-adenosylmethionine and catechol-O-methyltransferase in schizophrenia. *Am. J. Psychiat.* 128:1310–12
100. Dunner, D. L., Cohn, C. K., Weinshilboum, R. M., Wyatt, R. J. 1973. The activity of dopamine beta-hydroxylase and methionine-activating enzyme in blood of schizophrenic patients. *Biol. Psychiat.* 6:215–20
101. Pollin, W., Cardon, P. V. Jr., Kety, S. S. 1961. Effects of amino acid feedings in schizophrenic patients treated with iproniazid. *Science* 133:104–5
102. Brune, G. G., Himwich, H. E. 1962. Effects of methionine loading on the behavior of schizophrenic patients. *J. Nerv. Ment. Dis.* 134:447–50
103. Alexander, F., Curtis, G. C. III, Sprince, H., Crosley, A. P. Jr. 1963. L-methionine and L-tryptophan feedings in non-psychotic and schizophrenic patients with and without tranylcypromine. *J. Nerv. Ment. Dis.* 137:135–42
104. Park, L. C., Baldessarini, R. J., Kety, S. S. 1965. Methionine effects on chronic schizophrenics. *Arch. Gen. Psychiat.* 12:346–51
105. Haydu, G. G., Dhrymiotis, A., Korenyi, C., Goldschmidt, L. 1965. Effects of methionine and hydroxychloroquine in schizophrenia. *Am. J. Psychiat.* 122:560–64
106. Spaide, J., Tanimukai, H., Bueno, J. R., Himwich, H. E. 1968. Behavioral and biochemical alterations in schizophrenic patients. *Arch. Gen. Psychiat.* 18:658–65
107. Ban, T. A. 1969. On-going national collaborative studies in Canada: niacin in

the treatment of schizophrenia. *Psychopharmacol. Bull.* 5(3):5–20

108. Kakimoto, Y., Sano, I., Kanazawa, A., Tsujio, T., Kaneko, Z. 1967. Metabolic effects of methionine in schizophrenic patients pretreated with a monoamine oxidase inhibitor. *Nature* 216:1110–11
109. Antun, F. T. et al 1971. The effects of L-methionine (without MAOI) in schizophrenia. *J. Psychiat. Res.* 8:63–71
110. Gallagher, B. B. 1969. Amino acids and cerebral excitability. *J. Neurochem.* 16:701–6
111. Carson, N. A. J. et al 1963. Homocystinuria: a new inborn error of metabolism associated with mental deficiency. *Arch. Dis. Childhood* 38:425–36
112. Dunn, H. G., Perry, T. L. Dolman, C. L. 1966. Homocystinuria: a recently discovered cause of mental defect and cerebrovascular thrombosis. *Neurology* 16:407–20
113. Kaeser, A. C., Rodnight, R., Ellis, B. A. 1969. Psychiatric and biochemical aspects of a case of homocystinuria. *J. Neurol. Neurosurg. Psychiat.* 32:88–93
114. Spiro, H. R., Schimke, R. N., Welch, J. P. 1965. Schizophrenia in a patient with a defect in methionine metabolism. *J. Nerv. Ment. Dis.* 141:285–90
115. Frimpter, G. W., Haymovitz, A., Horwith, M. 1963. Cystathioninuria. *N. Engl. J. Med.* 268:333–39
116. Kline, N. S. et al 1967. Controlled evaluation of nicotinamide adenine dinucleotide in the treatment of chronic schizophrenic patients. *Brit. J. Psychiat.* 113:731–42
117. Ramsay, R. A., Ban, T. A., Lehmann, H. E., Saxena, B. M., Bennett, J. 1970. Nicotinic acid as adjuvant therapy in newly admitted schizophrenic patients. *Can. Med. Assoc. J.* 102:939–42
118. Ban, T. A., Lehmann, H. E. 1970. Nicotinic acid in the treatment of schizophrenias. *Can. Psychiat. Assoc. J.* 15:499–500
119. Mosher, L. R. 1970. Nicotinic acid side effects and toxicity: a review. *Am. J. Psychiat.* 126:1290–96
120. Meltzer, H., Shader, R., Grinspoon, L. 1969. The behavioral effects of nicotinamide adenine dinucleotide in chronic schizophrenia. *Psychopharmacologia* 15:144–52
121. Lipton, M. A. et al 1973. *Megavitamin and orthomolecular Therapy in Psychiatry.* Report of the APA Task Force on Vitamin Therapy in Psychiatry. Washington: Am. Psychiat. Assoc.
122. Baldessarini, R. J. 1966. Alterations in tissue levels of S-adenosylmethionine. *Biochem. Pharmacol.* 15:741–48
123. Mosher, L. R. 1970. Nicotinic acid side effects and toxicity: a review. *Am. J. Psychiat.* 126:1290–96
124. Wittenborn, J. R., Weber, E. S. P., Brown, M. 1973. Niacin in the long-term treatment of schizophrenia. *Arch. Gen. Psychiat.* 28:308–15
125. Pauling, L. Orthomolecular psychiatry. 1968. *Science* 160:265–71
126. Matthysse, S., Kety, S. S., Eds. 1974. See Ref. 68
127. Snyder, S. H., Banerjee, S. P., Yamamura, H. I., Greenberg, D. 1974. Drugs, neurotransmitters and schizophrenia. *Science* 184:1243–53
128. Matthysse, S. 1974. Implications of the catecholamine systems of the brain in schizophrenia. In *Brain Dysfunction in Metabolic Disorders,* ed. F. Plum. Res. Publ. Assoc. Nerv. Ment. Dis., 53:305–14. New York: Raven
129. Kebabian, J. W., Petzold, G. L., Greengard, P. 1972. Dopamine-sensitive adenylate cyclase in caudate nucleus of rat brain, and its similarity to the "dopamine receptor." *Proc. Nat. Acad. Sci. USA* 69:2145–49
130. York, D. H. 1972. Dopamine receptor blockade—a central action of chlorpromazine on striatal neurons. *Brain Res.* 37:91–99
131. Carlsson, A., Lindquist, M. 1963. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacol. Toxicol.* 30:140–44
132. Laverty, R., Sharman, D. F. 1965. Modification by drugs of the metabolism of 3,4-dihydroxyphenylethylamine, noradrenaline and 5-hydroxytryptamine in the brain. *Brit. J. Pharmacol.* 24:759–72
133. Nybäck, H., Borzecki, Z., Sedvall, G. 1968. Accumulation and disappearance of catecholamines formed from tyrosine-^{14}C in mouse brain; effect of some psychotropic drugs. *Eur. J. Pharmacol.* 4:395–403
134. Andén, N. -E., Butcher, S. G., Corrodi, H., Fuxe, K., Ungerstedt, U. 1970. Receptor activity and turnover of dopamine and noradrenaline after neuroleptics. *Eur. J. Pharmacol.* 11:303–14
135. Yoshida, M., Precht, W. 1971. Monosynaptic inhibition of neurons of the substantia nigra by caudato-nigral fibers. *Brain Res.* 32:225–28

136. Carlsson, A. 1975. Antipsychotic drugs and catecholamine synapses. See Ref. 68, pp. 57–64
137. Bunney, B. S., Walters, J. R., Roth, R. H., Aghajanian, G. K. 1973. Dopaminergic neurons: effect of antipsychotic drugs and amphetamine on single cell activity. *J. Pharmacol. Exp. Ther.* 185:560–71
138. Yaryura-Tobias, J. A., Diamond, B., Merlis, S. 1970. The action of L-dopa on schizophrenic patients (a preliminary report). *Curr. Ther. Res. Clin. Exp.* 12:528–31
139. Singh, M. M., Smith, J. M. 1973. Reversal of some therapeutic effects of an antipsychotic agent by an antiparkinsonism drug. *J. Nerv. Ment. Dis.* 157:50–58
140. Carlsson, A., Persson, T., Roos., B. -E., Walinder, J. 1972. Potentiation of phenothiazines by alpha-methyltyrosine in treatment of chronic schizophrenia. *J. Neural Transm.* 33:83–90
141. Janowsky, D. S., El-Yousef, M. K., Davis, J. M., Sekerke, J. 1973. Provocation of schizophrenic symptoms by intravenous administration of methylphenidate. *Arch. Gen. Psychiat.* 28:185–91
142. Angrist, B. M., Gershon, S. 1970. The phenomenology of experimentally induced amphetamine psychosis—preliminary observations. *Biol. Psychiat.* 2:95–107
143. Ellinwood, E. H. Jr., Sudilovsky, A., Nelson, L. M. 1973. Evolving behavior in the clinical and experimental amphetamine (model) psychosis. *Am. J. Psychiat.* 130:1088–93
144. Janowski, D. S., El-Yousef, M. D., Davis, J. M. 1973. Antagonistic effects of physostigmine and methylphenidate in man. *Am. J. Psychiat.* 130:1370–76
145. Cole, J. O., Clyde, D. J. 1961. Extrapyramidal side effects and clinical response to the phenothiazines. *Rev. Can. Biol.* 20:565–74
146. Schelkunov, E. L. 1967. Integrated effect of psychotropic drugs on the balance of cholino-, adreno-, and serotoninergic processes in the brain as a basis of their gross behavioral and therapeutic actions. *Activ. Nerv. Super.* 9:207–17
147. Snyder, S., Greenberg, D., Yamamura, H. I. 1975. Anti-schizophrenic drugs: affinity for muscarinic cholinergic receptor sites in the brain predicts extrapyramidal side effects. See Ref. 68 pp. 91–95
148. Bunney, S. Personal communication
149. Andén, N. -E. 1972. Dopamine turnover in the corpus striatum and the limbic system after treatment with neuroleptic and anti-acetylcholine drugs. *J. Pharm. Pharmacol.* 24:905–6
150. Andén, N. -E., Stock, G. 1973. Effect of clozapine on the turnover of dopamine in the corpus striatum and in the limbic system. *J. Pharm.* 25:346–48
151. Glowinski, J. 1975. New developments for the study of the mechanism of action of neuroleptics and amphetamine. See Ref. 68, pp. 81–85
152. Asper, H. et al 1973. Tolerance phenomena with neuroleptics: catalepsy, apomorphine stereotypes and striatal dopamine metabolism in the rat after single and repeated administration of loxapine and haloperidol. *Eur. J. Pharmacol.* 22:287–94
153. Puri, S. K., Lal, H. 1974. Tolerance to the behavioral and neurochemical effects of haloperidol and morphine in rats chronically treated with morphine or haloperidol. *Naunyn-Schmiedebergs Arch. Pharmakol. Exp. Pathol.* 282:155–70
154. Persson, T., Roos, B. -E. 1969. Acid metabolites from monoamines in CSF of chronic schizophrenics. *Brit. J. Psychiat.* 115:95–98
155. Rimon, R., Roos, B. -E., Rakkolainen, V., Alanen, Y. 1971. The content of 5HIAA and HVA in the CSF of patients with acute schizophrenia. *J. Psychosom. Res.* 15:375–78
156. Bowers, M. B. Jr. 1973. 5-Hydroxyindoleacetic acid (5HIAA) and homovanillic acid (HVA) following probenecid in acute psychotic patients treated with phenothiazines. *Psychopharmacologia* 28:309–18
157. Meltzer, H. Y., Sachar, E. J., Frantz, A. G. 1974. Serum prolactin levels in unmedicated schizophrenic patients. *Arch. Gen. Psychiat.* 31:564–69
158. Wise, C. D., Baden, M. M., Stein, L. 1975. Post-mortem measurement of enzymes in human brain: evidence of a central noradrenergic deficit in schizophrenia. See Ref. 68, pp. 185–98
159. Wyatt, R. J., Schwartz, M. A., Erdelyi, E., Barchas, J. D. 1975. Dopamine B-hydroxylase activity in brains of chronic schizophrenic patients. *Science* 187:368–70
160. Goldstein, M., Freedman, L. S., Ebstein, R. P., Park, D. H. 1975. Studies on dopamine-beta-hydroxylase in mental disorders. See Ref. 68, pp. 205–10

Copyright 1975. All rights reserved

THE ROLE OF RADIATION THERAPY IN THE DEFINITIVE TREATMENT OF ADENOCARCINOMA OF THE PROSTATE[1]

❖7173

Gordon R. Ray, M.D.[2] *and Malcolm A. Bagshaw, M.D.*
Division of Radiation Therapy, Department of Radiology, Stanford University School of Medicine, Stanford, California 94305.

INTRODUCTION

The use of ionizing radiation, both interstitial and external beam, in the definitive treatment of adenocarcinoma of the prostate has greatly increased during the last decade. This resulted from the realization that radical prostatectomy as a curative procedure is only applicable in 5–7% of patients (1), that endocrine intervention is associated with inherent limitations and potential morbidity (2), and from technological advances in the fields of radiation therapy and nuclear medicine. Although the first report of the use of megavoltage irradiation in the treatment of carcinoma of the prostate was published in 1964 (3), its exact role has not been clearly defined because of the paucity of relevant clinical studies. The purpose of this article is to present an overview of the pertinent information that has accumulated. It will include a brief discussion of the historical evolution of the use of ionizing radiation in the treatment of carcinoma of the prostate, a review of the results of treatment, emphasizing external beam megavoltage radiotherapy, and consideration of the role of radioisotopes in combination with various surgical procedures. In addition, preliminary data dealing with the use of external beam irradiation in patients with

---

[1] This work was supported in part by Grants CA-05838 and CA-15455 from the National Cancer Institute, National Institutes of Health, and the National Prostatic Cancer Project.

[2] Also affiliated with the Division of Radiation Therapy, Palo Alto Medical Clinic, Palo Alto, California 94301.

postprostatectomy residual or recurrent disease will be presented. The role of irradiation as a palliative measure will not be discussed.

General guidelines are presented as to patient selection, pretreatment evaluation, and radiotherapeutic techniques. The need for earlier detection of the tumor by the performance of routine rectal examinations and the establishment of well controlled, randomized clinical trials is stressed.

HISTORICAL EVOLUTION

The first radiotherapeutic approach to the management of carcinoma of the prostate was described in 1911 by Pasteau (4), who used an intraurethral radium source. Young (5) and Barringer (6) described modifications of this approach in 1917. Deming (7), expanding on Young's original experience, reported detailed results in 100 cases, noting a favorable response in 75%. The reports of Bugbee (8) and Bumpus (9) using intracavitary and interstitial irradiation provided further encouraging results. The ability of ionizing radiation to sterilize prostatic adenocarcinoma was confirmed in 1942 when Barringer (10) reported two patients treated by interstitial implantation in whom no tumor was found at necropsy 6 and 7 years respectively after treatment.

Reports (11, 12) of external beam therapy during the period 1930 to 1955 dealt mainly with results of treatment with kilovoltage (250 KVP) radiation. Inherent with this quality of radiation was the deposition of maximum energy in the epidermis and excessive scattering of radiation which prohibited delivery of a tumoricidal dose to the prostate without unacceptable damage to the skin and normal tissue surrounding the prostate. Even with these limitations, Smith (13) and Widmann (14) both reported good results in the relief of pain associated with bony metastases, and the latter presented impressive results in 82 cases treated for palliation of symptoms associated with advanced local neoplasm. All forms of treatment were eclipsed in 1941 when Huggins, Stevens & Hodges (15) showed that the growth of prostatic carcinoma could be retarded by androgen deprivation. However, as the limitations of hormonal treatment became apparent later, interest renewed in the radiotherapeutic management of prostatic carcinoma. In 1952, Flocks et al (16) showed that the infiltration of a radioactive gold (^{198}Au) colloidal suspension directly into the neoplasm and/or periprostatic soft tissues following subtotal resection decreased the local recurrence rate.

The use of ionizing radiation, stimulated by the report of Flocks, prompted a review of the possible use of external beam radiotherapy. The latter was in part a result of the availability of new megavoltage ($>$ 1 MeV) radiation therapy units which allowed the delivery of a tumoricidal dose to the prostate without excessive damage to the skin or other normal tissues. In 1962, Bagshaw & Kaplan (17) presented their preliminary observations on the application of external beam irradiation as a definitive treatment modality in carcinoma of the prostate at the Tenth International Congress of Radiology. The first published report appeared in 1964 when Budhraja and Anderson presented the results of treating 36 patients using either 2 MeV X rays or a ^{60}Cobalt teletherapy unit. This was followed in the next five years by other reports which attested to the radiosensitivity of carcinoma of the prostate and the feasibility of definitive treatment by external beam megavoltage irradiation.

EXTERNAL BEAM MEGAVOLTAGE RADIATION THERAPY

The following section presents the results of treatment of 430 patients referred to the Division of Radiation Therapy, Stanford University School of Medicine for definitive radiation therapy, and reviews pertinent data of other investigators. Patients were referred to the Division of Radiation Therapy during the period September 1956 to December 31, 1972 and were followed through June 30, 1974 for periods ranging from 1.5 to 16 yr. Mean follow-up was 6.4 yr. Patients were included in the study if histological confirmation was present, appropriate studies had ruled out the presence of metastases, and if no prior definitive surgery had been performed. The majority of these patients have been included in previous reports (18, 19); however, follow-up has been extended and additional patients accessioned. All patients were treated with either 4.8 or 4 MeV X rays produced by a linear accelerator, with the intent to irradiate homogeneously a small volume of tissue including the prostate, periprostatic soft tissue, bladder neck, and seminal vesicles. An isocentric (20), rotational technique was used to deliver a dose of 7000–7500 rads (at the 100% isodose contour) through field sizes ranging from 6.6 X 6.6 cm to 8.8 X 8.8 cm. The

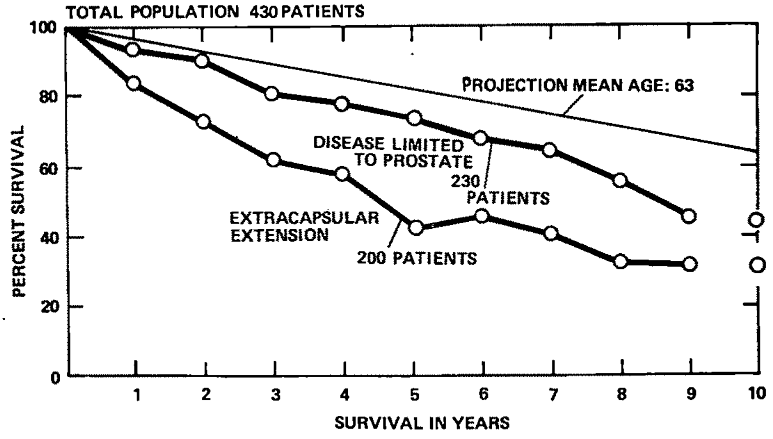


Figure 1 Actuarial survival curves for patients with prostatic carcinoma; comparison by local extent as determined by digital rectal examination. Survival curves are uncorrected; the projected survival for a population of normal males with a mean age of 63 is provided for comparison only. The survival rate of the DLP group is significantly ($p < 0.05$) better than that of the ECE group.

Table 1 5- and 10-year uncorrected actuarial survival rate (% survival) according to clinical stage

Clinical Group	Total Number of Patients	Actuarial Survival Rate 5-yr	10-yr
Disease Limited to Prostate[a] (Stage B[b], Stage II[c])	230	71	45
Extracapsular Extension[a] (Stage C[b], Stage III[c])	200	41	31

[a] Stage does not change with elevated serum acid phosphatase level.

[b] American Urological System; stage does not change with elevated serum acid phosphatase level.

[c] V.A. Cooperative Study Group; patients must have normal serum acid phosphatase or else are classified as Stage IV irregardless of extent of local disease.

details of tumor localization and treatment techniques have been published previously (21).

Patients were divided into two clinical groups based upon the extent of disease as determined by digital rectal examination. Two hundred thirty patients had disease limited to the prostate (DLP group) and 200 patients had extracapsular extension into the seminal vesicles and/or other periprostatic tissues. The latter were termed the extracapsular extension group (ECE group). Survival data (calculated as beginning with the first treatment date) were analyzed according to the actuarial or life table method (22) for these two groups (Figure 1), and the direct method. Table 1 shows the 5- and 10-yr actuarial survival rates for the DLP and ECE groups, respectively. The clinical states analogous to these groups using the criteria of the American Urological System (23) and those employed by the Veterans Administration Cooperative Study Group (24) are also provided. The status and survivorship (calculated according to the direct method) of all patients at risk for 5 and 10 yr is shown in Table 2. For the purpose of this analysis, patients were considered to have died without cancer if they were examined 6 months or less prior to death and found to be free of malignancy, and if they died of an unrelated cause found at autopsy or listed on the death certificate. The significant number of patients alive without clinical evidence of disease should be noted as well as the percentage of patients dying free of cancer. The latter patients are considered treatment failures and to have died of cancer in the calculation of survival by the actuarial method.

The direct method differs from the actuarial method in that the latter describes a survival *pattern* rather than a survival rate at one specific end point (i.e. 5 and 10 yr). Furthermore, the actuarial method utilizes survival information accumulated on *all* patients up to the closing date of the study, and not information limited to those patients actually observed for the specified period as in the direct method.

Although it was our policy to withhold hormonal manipulation until the development of symptomatic local recurrence or metastatic disease, approximately 40% of

Table 2 Survival status of patients at risk 5 and 10 years (direct survival)[a]

Group[b]	No. of Patients	No. Alive Without Cancer (%)	No. Alive With Cancer (%)	No. Dead Without Cancer (%)	No. Dead With Cancer (%)
Minimum of 5 yr					
DLP	86	60(70)	3(3)	12(14)	11(13)
ECE	88	33(37)	8(9)	11(13)	36(41)
Minimum of 10 yr					
DLP	31	13(42)	2(6)	11(35)	5(16)
ECE	28	8(29)	1(3)	5(18)	14(50)

[a] The direct survival method uses follow-up information limited to those patients actually at risk for the specified period compared to the actuarial method which utilizes follow-up information accumulated on *all* patients up to the closing date of the study.

[b] DLP = disease limited to prostate; ECE = extracapsular extension. Numbers in parentheses represent the percentage of the total number of patients in each group.

the patients had already been castrated and/or received estrogens prior to referral. A comparison of the 5- and 10-yr direct survival rates between the DLP and ECE groups who did and did not receive preradiotherapy hormonal manipulation revealed no significant difference (Table 3). Therefore, we concluded that hormonal manipulation did not favorably bias our long term survival results.

Table 4 lists in chronological order the pertinent reports of other investigators dealing with the definitive treatment of adenocarcinoma of the prostate with megavoltage irradiation. These all suffer to a certain degree, as does the Stanford experience, from small sample size, inadequate controls, lack of 10- or 15-yr follow-up, and survival analysis that too often does not distinguish between patients treated by irradiation alone and those having concomitant hormonal manipulation. Furthermore, there has been no definitive analysis of the relation between histologic grade and stage, response to radiotherapy and/or survival. These series should not be considered comparable because of the various selection biases, the different treatment techniques utilized, and the wide spectrum in total dose delivered to the prostate, the time in which it was delivered, and the volume treated (time-dose-volume relationship).

Budhraja & Anderson (3) in 1964 reported a group of 53 patients, 36 of whom were treated by megavoltage irradiation with concomitant hormone therapy. The 3- and 5-yr survival rates were 60 and 25% respectively. Bagshaw, Kaplan & Sagerman (25) in 1965 reported a 60% 5-yr actuarial survival rate in 65 patients treated with 4.8-MeV X rays. In 1967, Del Regato (26), utilizing a three-field technique and a dose of 7500–8000 in 5-6 weeks, reported on 7 patients, 4 of whom were alive and well 5 yr after receiving irradiation for "locally inoperable" carcinoma of the prostate. Two of the 4 patients had received concurrent hormonal

Table 3 Survival status of patients with and without preradiotherapy hormonal manipulation at risk 5 and 10 years (direct survival)

Group[a]	Total[b] No. of Patients		No. Alive Without Cancer (%)		No. Alive With Cancer (%)		No. Dead Without Cancer (%)		No. Dead With Cancer (%)	
	5 yr	10 yr	5 yr	10 yr	5 yr	10 yr	5 yr	10 yr	5 yr	10 yr
With Preradiotherapy Hormonal Intervention										
DLP	19	9	13(68)	2(22)	1(5)	—	2(11)	4(44)	3(16)	3(33)
ECE	38	21	12(31)	6(29)	3(8)	1(5)	6(16)	3(14)	17(45)	11(52)
Without Preradiotherapy Hormonal Intervention										
DLP	58	20	40(69)[c]	11(55)[c]	2(3)	2(10)	8(14)	5(25)	8(14)	2(10)
ECE	39	7	13(33)[c]	2(29)[c]	4(10)	—	5(13)	2(29)	17(44)	3(42)

[a] DLP = disease limited to prostate; ECE = extracapsular extension.

[b] Total No. of Patients at risk in each group is less than those at risk when considering total population (Table 2), since patients had to be excluded from this analysis if information regarding hormonal status was lacking.

[c] None of the patients alive without cancer received supplemental hormonal manipulation.

therapy. In 1968 Dykhuizen et al (27) provided a follow-up of cases first reported by George et al (28). Patients were treated with ^{60}Cobalt and received a dose of 6000–75000 rads in 10–11 weeks. Of 8 patients with intracapsular disease (Stage B) treated by irradiation alone, 7 were alive without evidence of tumor 1 to 4.5 yr following irradiation. All 7 had a postirradiation biopsy of the prostate which was negative for tumor. The 8th patient died of metastatic tumor with local control. Fourteen patients with extracapsular disease (Stage C) were also treated. Nine of the 14 patients had local control of their neoplasm for periods ranging from 1 to 6.5 yrs following irradiation. Two of them had estrogens and orchiectomy prior to radiation. Eight of the 9 patients had negative needle biopsies of the prostate following irradiation. Bennett (29), using a four-field technique and ^{60}Cobalt, reported encouraging results following treatment of 10 patients who were followed from 1 to 80 months. Grout et al (30) using ^{60}Cobalt and a three-field technique noted that 39 of 41 patients who received doses of between 5300–7200 rads were alive from 3–60 months since "diagnosis" without hormonal supplementation. Thirty five of these patients had extracapsular extension (Stage C). Twenty five of the 41 patients were alive without clinical evidence of tumor, but 3 had persistent tumor on repeat biopsy. Although these results are promising, only 11 patients were followed for more than 24 months. Odell et al (31) reported on 68 patients treated to the prostate with ^{60}Cobalt. Tumor doses ranged from 4000 to 6800 rads delivered in 5–7 weeks. A 5-yr actuarial survival rate of 60% was noted for a subgroup of 47 patients. Unfortunately, clinical staging was not reported, a large proportion of the entire group had some form of hormonal manipulation, and we are not told the exact number of patients followed for 5 and 10 yr among the subgroup of 47 patients. Carlton et al (32) reported the results of treatment in 106 patients treated with curative intent. These patients received a total dose to the prostate of 6000–7500 rads delivered by a combination of interstitial irradiation using radioactive gold seeds followed by external irradiation. The exact time-dose-volume relationships were not provided. Of those patients with Stages, A, B, and C[1]$_1$ tumors treated by irradiation alone and followed for periods ranging from 1–8 yr, 3 died. However, only a small percentage were followed for periods greater than 4 yr, and the number of patients alive without clinical evidence of disease was not indicated. Loh, Brown & Beiler (33) treated 36 patients; 7 had Stage B, 28 Stage C, and 1 Stage D prostatic cancer. Follow-up extended from 5 months to 5 yr posttreatment. Patients with Stage B tumors were treated to a small volume using field sizes that ranged from 8 X 8 cm to 10 X 10 cm. Patients with Stage C lesions were treated with a technique which included the regional nodes to a dose of 5000 rads followed by a prostatic supplement of 2000 rads. Total dose to the prostate ranged from 6000 to 7580 rads delivered in 6 to 7 weeks. Six of the 7 patients with Stage B disease were alive without clinical or biopsy proof of local disease. The seventh patient died of pneumonia without clinical evidence of cancer. Of twenty eight patients with Stage C lesions, 16 (57%) were free of cancer. The investigators have also correlated their results of treatment with stage and grade of tumor. Of the 7 patients with Stage B disease, 6 had lesions of intermediate grade and 1 was of high grade. All patients were free of cancer. The 28 patients with Stage C lesions consisted of 15 with intermediate grade malignancies and 13 whose tumors were less well differentiated. Thirteen (87%) of the former "responded" compared to only 3 of the 13 patients with less well differentiated tumors. More detailed information on this same group of patients is reported in a more recent report (34). Hazra (35) reported a 60% 5-yr actuarial survival rate in 30 patients, 25 of whom were referred after failure of hormonal manipulation. The majority of patients received 5500 rads in 4 weeks using a three-field technique. Grossman et al (36), reporting on 58 patients treated to doses of 7200 rads in 11 weeks and followed from 12–24 months, showed the apparent feasibility of treating extensive volumes without undue patient morbidity. The data of investigators using a variety of techniques and sources of radiation (Table 4) tend to support our impression that potentially tumoricidal doses of irradiation can be delivered to the prostate with relative safety. In general, initial response of the local tumor to irradiation and survival rates at 5 yr (when provided) were encouraging.

RADIATION RESPONSIVENESS

In our experience, a change in size and/or hardness of the prostate occurred in only 30% of patients by the completion of irradiation; but an objective response was noted in 80% by 6 months. Serial rectal examinations after 6 months elapsed revealed one of four courses: the most common was a significant decrease in size

---

[1] C$_1$ is a special subcategory of Stage C and consists of patients with tumors that extend into the periprostatic area but are smaller than 6 cm in diameter.

Table 4 Definitive radiotherapeutic management of carcinoma of the prostate (published series)

Author	Total No. Pts.	Treatment Volume	Technique	Total Dose To Prostate (Rads/Wks)	Follow-up (months)	Radiation Source	Ref.
Budhraja & Anderson (1964)	36	NR	3-Field	$\frac{5000}{5}$	1–60	^{60}Co, 2 MeV X-Rays	3
George et al (1965)	9	R	C-PO	$\frac{6000\text{–}7500}{7\text{–}11}$	5–54	^{60}Co	28
Bagshaw et al (1965)	65	L	ROT	$\frac{7000\text{–}7500}{7\text{–}8}$	12–96	4.8 MeV-LA	25
Del Regato (1967)	7	R	3-Field	$\frac{7500\text{–}8000}{6}$	24–60	^{60}Co	26
Dykhuizen et al (1968)	22	R	C-PO	$\frac{6000\text{–}7500}{10\text{–}11}$	4–54	^{60}Co	27
Bennett (1968)	10	L	4-Field	$\frac{6000}{6}$	1–80	^{60}Co	29
Grout, DC et al (1971)	41	R	3-Field	$\frac{5300\text{–}7200}{6\text{–}8}$	3–60	^{60}Co	30
Odell, A et al (1971)	68	R	3-Field	$\frac{4000\text{–}6800}{5\text{–}7}$	16–120	^{60}Co	31
Loh, ES et al (1971)[a]	7 29	L R	ROT 4-Field + ROT	$\frac{6000\text{–}7500}{6\text{–}7}$	5–60	^{60}Co	33

Table 4 (continued)

Rhamy, RK et al (1972)	15	L	ROT	$\frac{7000-7500}{7}$	NR–48	^{60}Co; 6 MeV-LA	38
Hazra, TA (1972)	30	R	3-Field	$\frac{4000-6200}{3-7}$	12–61	^{60}Co	35
Carlton, CE et al (1972)	106	NR	NR	$\frac{6000-7500}{NR}$	12–96	^{60}Co; B, LA	37
Phillips & Lattimer (1972)	33	R	3-Field	$\frac{5900-7100}{5-7}$	12–36	^{60}Co; B, LA	37
Ray, GR et al (1973)[b]	310	L	ROT	$\frac{7000-7500}{7-8}$	24–180	4.8 MeV-LA	18
Grossman et al (1974)	58	R;S-R#	C-PO;LAT	$\frac{7200}{11}$	12–24	B, 4 MeV-LA	36
Bagshaw et al (1974)	11	L	ROT	$\frac{7000-7500}{7-8}$	6–24	4 MeV-LA	43
	16	R	4-Field+ROT	$\frac{7000}{7-8}$			
	7	S-R	4-Field+ROT				

[a] Michaels, M. M. et al 1974 – Follow-up communication (34).

[b] Includes patients reported by Bagshaw, M. A. et al 1965 (25).

[c] Abbreviations: C-PO, Co-planar opposed; ROT, Rotation; LAT, Lateral; NR, Not recorded; L (Local), Limited to prostate and periprostatic tissues; R (Regional), Prostate + pelvic lymph nodes; S-R (Supra-Regional), Prostate + pelvic lymph nodes + periaortic nodes to level of T_{12}, # includes patients treated above T_{12}; LA, Linear accelerator; B, Betatron.

and then a stabilization of the gland size at 1 yr followed by progressive induration of a majority of the gland presumably due to radiation-induced fibrosis; 15% of glands returned to normal size and consistency; 20% showed progressive diminution of the gland until a flat or concave prostatic bed was palpated; and, lastly, progressive enlargement of the gland and/or the reappearance of nodularity occurred in 10–15% of cases. The latter was usually a sign of active tumor. Sterilization of local prostatic neoplasm was documented by serial step sectioning of the glands of 13 patients, 9 of whom died of intercurrent disease and 4 of distant metastases. Phillips & Lattimer (37) reported a change of size and/or hardness of the prostate in only 1 of 12 patients at the completion of irradiation, but 75% showed a significant response by 2 months, and 85% were of normal size and consistency at 9 months. However, 11 of the 12 patients had concomitant hormonal manipulation. Grout et al (30) noted regression of the local lesion in 27 of 37 patients that was rapid in some patients and took several months in others.

RADIATION FAILURE

Treatment failure was manifest as local persistence or recurrence of tumor in the prostate and/or distant metastases. In our series, documentation of failure was based primarily on physical examination and/or roentgenographic findings, biopsy confirmation being rarely obtained. Seventy percent of 120 failures were recorded within 24 months after irradiation; however, 5% occurred as late as 73 to 138 months. Of those patients who failed, 35% presented with local and disseminated disease, 60% with metastatic foci and apparent local control, and 9% as local failures only.

In our experience, local failure as determined by serial rectal examination occurred in 53 (12%) of 430 patients. However, reservations as to the validity of our criteria for local control may be raised by the recent description by other investigators of persistent carcinoma in biopsies of the prostate, some of normal consistency, months to years following irradiation (32, 38). A recent report by Cox & Tijerina (39) presented preliminary results of 68 transperineal prostatic biopsies obtained from 30 patients 3–30 months after completion of radiation therapy. They noted that 26 of 43 biopsies performed less than 9 months postirradiation were positive for tumor, while 20 of 25 biopsies performed 12–30 months postirradiation were negative. Unfortunately, 10 of their patients had orchiectomy prior to treatment, and the percentage of positive biopsies which subsequently became negative was not recorded. We have observed 2 patients with positive transrectal needle biopsies at 9 and 13 months postirradiation whose subsequent biopsies were negative at 24 months. A possible explanation for these findings is that since the cytological criteria of malignancy do not necessarily imply reproductive integrity, cells appearing malignant in postirradiation biopsies of the prostate may not have the potential for unlimited proliferation. We believe these results do attest to the slow regression rate of prostatic carcinoma and infer that a positive biopsy obtained prior to 15 months is often erroneous, since many will subsequently become negative. However, the real significance of a positive biopsy following irradiation of the prosate in asymptomatic patients in whom there is no evidence of tumor growth on rectal examination will have to await future reports of the fate of such patients who are not given adjuvant therapy, and which relate serial biopsy results to stage and time-dose-volume relationships.

SYMPTOMS AND COMPLICATIONS

The predominant symptoms and complications associated with the radiotherapeutic management of carcinoma of the prostate involve the gastrointestinal and/or genitourinary system. We considered symptoms acute if they ceased within 3 months of completion of radiation therapy and persistent if they continued beyond 3 months; mild if they did not cause interruption of radiotherapy and patient duress was not excessive, and severe if they did. In our experience, acute, mild symptoms were seen in 40% of patients. The most common acute symptoms were diarrhea, rectal urgency, dysuria, and frequency. Acute, severe symptoms referable to either the GI or GU tract were recorded in 5 and 4% of the patients respectively. Persistent, severe symptoms were recorded in 51 of 430 patients (12%). These consisted primarily of blood-streaked stools, tenesmus, rectal urgency, and dysuria. Although classified as persistent and severe, the majority did resolve within one year with conservative, symptomatic management.

In general, the results of other investigators support our impression that the majority of patients tolerate treatment well, acute symptoms are mild or moderate in severity, respond to conservative, symptomatic management, and usually disappear within 3–6 months of the completion of radiation therapy.

Serious complications encountered by other investigators appear to be relatively infrequent and limited to the GI or GU systems. In our experience, urethral stricture developed in 16 of 430 patients. The majority were associated with multiple predisposing factors and occurred during the early years of the study. Del Regato (26) noted urethral strictures only in those patients who had undergone prior transurethral resection of the prostate (TURP); however, Loh, Brown & Beiler (33) reported that 5 of 8 patients with urethral strictures following irradiation had not had a prior transurethral resection. We found no significant difference in the incidence of urethral strictures following irradiation between those patients having one TURP and those having none. However, multiple (>3) TURPs did seem to increase the incidence of postradiation urethral stricture. We recommend minimal instrumentation of the lower urinary tract prior to and following irradiation, and a 4–6 week waiting period following TURP before initiating irradiation. Although fistulization has not been a significant problem, Grout et al (30) reported 1 case of a vesicorectal fistula, and 1 patient in our series required a colostomy due to rectal ulceration. We have attributed this to radiation; however, viable tumor was also present. Although fecal and urinary stress incontinence were not noted in our previous reports, we have recently encountered 3 cases of the former and 2 of the latter. Grout et al (30) also noted 4 patients with stool and 1 with urinary stress incontinence.

IMPOTENCY

The incidence of impotency reported by other investigators following radiation therapy varies between 41 and 84%. We chose to take the presence or absence of sexual activity at 15 months postirradiation as the end point for analysis. Potency was maintained in 65 of 110 patients (59%) (mean age 59) who were potent prior to irradiation and did not receive posttreatment endocrine intervention. Rhamy, Wilson & Caldwell (38) reported that 8 of 15 patients (53%) whose mean age was 57.4 were potent following irradiation. However, Loh, Brown & Beiler (33) noted that, following irradiation, 16 of 19 patients (84%) were impotent. Neither the mean age of the patients nor the time between irradiation and analysis was given. The wide variability in incidence may be a reflection of different treatment techniques, the veracity of the patient, or the tenacity of the individual obtaining the history.

RADIOISOTOPES

Radioisotopes have been used in the treatment of prostatic carcinoma alone or in combination with surgical procedures to achieve cure and/or relief of local symptoms. In addition, intravenously administered radioactive phosphorus (^{32}P) is used for palliation of bone pain (40, 41).

The use of radium by early investigators has already been alluded to. Subsequently, isotopes with lower mean energies or with beta emission were introduced which were more suitable for irradiation of small volumes homogenously without unnecessary irradiation to normal adjacent tissues. Flocks used local infiltration of a colloidal suspension of radioactive gold (^{198}Au) to successfully reduce the incidence of postprostatectomy recurrences, and also produce encouraging 10-yr survival rates in selected patients. Other investigators have combined interstitial implantation of the prostate with pelvic lymphadenectomy or external irradiation in an attempt to cure selected patients with Stage B and C disease. The work of Carlton et al (32) using radioactive gold seeds plus supplemental external irradiation to the prostate and regional nodes was discussed. Whitmore et al (42) recently described a technique employing retropubic implantation of the prostate with ^{125}I seeds and concurrent pelvic lymphadenectomy. The total dose to the prostate from the permanent implant was 8000 rads at 2 months and 16,000 rads at 1 yr. Although short term follow-up precluded meaningful survival analysis, 17 of 26 patients (65%) followed from 3 to 24 months were felt to show local regression. This approach has the advantage of more accurate staging, definitive treatment requiring only one hospitalization, and minimal GI and GU morbidity.

PROSTATIC CARCINOMA AND LYMPHATIC METASTASES: IMPLICATIONS FOR THERAPY

Carcinoma of the prostate may disseminate by either the vascular and/or lymphatic routes. Therefore, its potential curability depends first upon the exclusion of patients with hematogenous metastases and then the ablation of the primary tumor as well

Table 5 Incidence of lymph node metastases related to clinical stage

	Disease Limited to Prostate (Stage B, II)			Extracapsular Extension (Stage C, III)			
Author	Total No. Patients	No. with Positive Nodes	%	Total No. Patients	No. with Positive Nodes	%	Ref.
Whitmore & MacKenzie (1959)[a]	2	0	0	18	9	50	48
Flocks et al (1959)[a]	29	2	7	384	144	38	44
Arduino & Glucksman (1962)[a]	44	14[d]	32	—	—	—	46
Whitmore et al (1972)[a]	8	2	25	18	10	55	42
McCullough et al (1974)[a]	4	1	25	35	18	51	45
Bagshaw et al (1974)[b]	18	2	11	16	11	69	43
Ray et al (1974)[b]	11	4	36	15	8	53	c

[a] Data from pelvic lymphadenectomy or exenteration.

[b] Data from transperitoneal diagnostic laparotomy and selective node biopsy.

[c] In preparation.

[d] Ten of eleven (91%) patients with unsuspected microscopic seminal vesicle invasion had pelvic lymph node metastases.

as its lymphatic extensions. It is important to emphasize that there appears to be a significant number of patients in whom lymphatic dissemination occurs without clinically detectable extranodal (osseus and/or soft tissue) distant metastases. Table 5 summarizes pertinent data on the incidence of lymph node metastases as related to clinical stage. Prospective data on the relative incidence of metastases to the presumed primary (obturator, hypogastric) and secondary (common, external iliac, and periaortic) echelon nodes are scant. In a selected group of patients, we have found that all patients with metastases to the periaortic nodes also had pelvic lymph node involvement.[3] Bagshaw et al (43) recently reported preliminary data in a group of 34 consecutive patients who had transperitoneal exploration and selective node biopsy within 2 weeks of lymphangiography. Of 18 patients with Stage B disease, 2 were found to have positive nodes. One of the 2 had periaortic involvement. Eight of 16 patients with Stage C disease had positive nodes at exploration, 2 with involvement of the pelvic nodes alone, 5 with pelvic and periaortic involvement, and 1 patient with a positive periaortic node and negative pelvic nodes. Bone scans using ^{99m}Tc-labeled diphosphonate, bone marrow aspirations, and biopsies were negative in all patients.

It is evident that lymphatic dissemination may occur early in the period of manifest clinical disease (Table 5) and the incidence increases with size or dedifferentiation of the primary tumor (44, 45) and/or invasion of the seminal vesicles (1, 46). Similar factors are also associated with decreased survival rates following radical perineal prostatectomy. Although a 10-yr survival rate of 33% can be attained with radical prostatectomy, the severity of the selection process limits this operation as a curative procedure to less that 5% of patients (47). The value of more extensive surgery is in doubt. Whitmore & MacKenzie (48) reported that 0 of 9 patients with regional lymph node metastases survived 5 yr following radical cystectomy or exenteration.

---

[3] Ray, G. R. et al. *Operative Staging of Apparently Localized Prostatic Carcinoma* (Unpublished).

The significant incidence of lymph node metastases in patients with carcinoma of the prostate and the limited number of patients that could be cured by surgery alone prompted the development of new diagnostic and therapeutic approaches and modification of old ones. New therapeutic approaches involved the combination of lymphadenectomy with various forms of irradiation. In actuality, this approach was proposed by Flocks (16) nearly two decades previously. In addition to Flocks' use of a radioactive gold colloidal suspension, interstitial implantation of the prostate with or without the use of external irradiation and concomitant pelvic lymphadenectomy was also advocated as a new therapeutic approach. Modifications of existing approaches involved continued reliance on external beam irradiation, but with the addition of diagnostic and therapeutic techniques which attempted to more precisely identify and treat the primary tumor and its lymphatic extensions. Bipedal lymphangiography, preferably confirmed by selective node biopsy, was used to determine the extent of lymphatic involvement (49). An attempt to tailor the radiation portals to include all known areas of disease resulted in three therapeutic strategies which we have designated as local, regional, and supraregional. Local treatment refers to irradiation of the prostate and immediate periprostatic tissues, regional treatment to irradiation of a volume extending to the level of the bifurcation of the common iliac artery or aorta, and supraregional to treatment of the periaortic nodes in addition to the volume included under regional treatment. We have not treated prophylactically above the level of the diaphragm; however, others (36) have advocated treatment of the mediastinum and bilateral supraclavicular nodes in selected patients. A limited number of patients and short term follow-up precluded meaningful survival analysis of our patients who received regional or supraregional treatment. We must stress that the ability of irradiation to favorably alter the natural history of prostrate carcinoma once periaortic metastases have occurred is pure conjecture at this time. Limitations are inherent with either the new approaches or the modified old ones. Reliance on pelvic lymphadenectomy alone offers little or no chance of cure in patients with involvement of the periaortic nodes. Similarly, since the interpretation of nodes opacified by lymphangiography is associated with a 20% false negative rate and the procedure does not opacify primary echelon nodes, reliance on this procedure alone may result in understaging the tumor and thereby lead to an inappropriate treatment strategy. It is imperative, therefore, that criteria be established for selection of patients who will benefit most from a particular approach. Ultimately, a highly individualized approach relying on a combination of these procedures with the possible addition of systemic chemotherapy or hormonal intervention may prove most successful.

POSTPROSTATECTOMY RESIDUAL OR RECURRENT TUMOR

A significant number of patients undergoing radical retropubic or perineal prostatectomy have incomplete excision of their tumor or subsequently develop local recurrence. The potential role of external beam irradiation in such patients has not been explored extensively; however, we have recently presented preliminary observations on the results of treatment of 21 such patients (50). Eleven patients were referred for irradiation because of incomplete excision and 10 for palpable local recurrence. Eight of 21 patients were alive without clinical evidence of disease for periods ranging from 29–97 months after radiation therapy. The best survival was in the 11 patients referred because of incomplete excision. Nine were alive, 6 without cancer, while 2 were dead of intercurrent disease without evidence of cancer. Two of 10 patients referred with palpable local recurrence were alive without neoplasm. All patients alive without cancer were treated by irradiation alone without hormonal intervention.

GUIDELINES

The following sections suggest general guidelines for the selection of patients most suitable for definitive treatment by irradiation and their pretreatment evaluation. Treatment techniques presently used by the Division of Radiation Therapy at Stanford to administer external beam irradiation are also presented.

Patient Selection

Patients with tumors that are well differentiated and no larger than 1 cm in diameter at rectal examination are best treated by radical prostatectomy unless precluded by medical contraindications or the patient's fear of impotency. Earlier detection by the performance of routine, meticulous rectal examinations on asymptomatic patients is mandatory if an increased number of patients are to be cured by either radiation therapy or radical prostatectomy. The importance of earlier detection and treatment is attested to by the better survival rates achieved by radiation therapy in the DLP group as compared to the ECE group (Table 1) and by the excellent results obtained with radical prostatectomy in a highly selected group of patients reported by Gilbertson (51). He noted respective 5- and 10-yr absolute survival rates of 91 and 69% in 23 of 75 patients undergoing radical prostatectomy following the detection of asymptomatic nodules at rectal examination at a cancer detection center.

Since only 5% of patients with prostatic carcinoma are amenable to radical prostatectomy and 50% present with metastases, there remains 45% of patients in whom no definitive form of therapy has been available previously. This latter group of patients, whose tumors are apparently localized to the prostate but too extensive for radical prostatectomy, or have invaded the seminal vesicles or periprostatic soft tissues, should be considered candidates for definitive radiation therapy. We believe external beam megavoltage irradiation to be the most widely applicable, potentially curative of the various therapeutic modalities available in these patients. Other approaches using ionizing radiation as part of an overall treatment program (i.e. lymphadenectomy plus external and/or interstitial irradiation of the prostate) offer additional flexibility. We feel that the criteria for selection of patients for these latter approaches are best established by the original investigators, since the data presently available are only preliminary in nature.

The most appropriate volume to be irradiated in each patient receiving external beam radiotherapy can only be determined after a thorough pretreatment evaluation establishes the extent of the disease. Our recommendations for this evaluation and the therapeutic implications of the information obtained are presented in the following section.

PreTreatment Evaluation and Treatment Volume

It is our belief that all patients with apparently localized prostatic carcinoma should have the following evaluation prior to initiating definitive radiation therapy: (*a*) complete history and physical examination, (*b*) total and tartrate-inhibited serum acid phosphatase obtained prior to digital examination, (*c*) chest roentgenogram, (*d*) intravenous pyelogram, (*e*) ^{99m}Tc-labeled diphosphonate bone scan, (*f*) percutaneous posterior iliac crest Jamshidi needle biopsy, and (*g*) bipedal lymphangiography.

The significant incidence of lymph node metastases in apparently early carcinoma and the 20% false negative rate associated with lymphangiography make clinical staging imprecise. Other investigators have stated that operative staging is mandatory (32). We found that lymphangiography followed by transperitoneal laparotomy and selective node biopsies revealed that approximately 15–25% of patients with Stage B tumors and 50–60% of patients with Stage C tumors were understaged (i.e. had pelvic and/or periaortic lymph node metastases). It remains to be determined whether operative staging should be a routine procedure and whether it should be performed by a transperitoneal or an extraperitoneal approach. Our initial experience with transperitoneal exploration followed by high dose megavoltage irradiation revealed a relatively high incidence of small bowel complications. This has, however, apparently been greatly reduced by recent changes in radiotherapeutic techniques. We hope, as the result of ongoing trials, that a prognostic index based on multiple parameters, i.e. local extent of tumor, histology, serum and bone marrow acid phosphatase, and lymphangiography, will allow the prediction of patients most likely to have lymph node and/or vascualr dissemination. Therefore, exploration will be obviated in this group of patients and the development of therapeutic policies facilitated. Until that time, we recommend the following general guidelines and operational schema (Figures 2 and 3) be used. Two major categories of patients are considered; the first consists of patients whose clinical stage is confirmed by surgical exploration, and the second of patients staged by clinical means alone.

OPERATIVE STAGING

Local Treatment

Patients with a negative lymphangiogram whose surgical exploration reveals the obturator, hypogastric, and external iliac lymph nodes to be free of tumor can be treated by radiation portals equal to or less than 8 X 8 cm. Such treatment will include the prostate, seminal vesicles, periprostatic soft tissues, and immediate periprostatic nodes.

Regional and Supraregional Treatment

This approach is reserved for those patients in whom the pelvic and/or periaortic nodes are proven to contain tumor. Prophylactic treatment of the periaortic nodes when pelvic nodes are positive is not recommended unless performed on an investigational basis. Patients in whom only the hypogastric and/or obturator nodes contain tumor may be considered candidates for a modified regional approach where the superior level of the treatment volume is at the bifurcation of the common iliac artery, and not the aorta as in the regional approach.

CLINICAL STAGING

Figures 2 and 3 depict operational schema to be followed when the stage of disease is determined by clinical methods alone. They reflect an appreciation of the limitations inherent in clinical staging, and attempt to improve its precision by relying on various parameters known to be associated with an increased probability of lymphatic dissemination. These schema were based on the following: (*a*) that rectal examination often underestimates the extent of local neoplasm (47, 52); (*b*) that the

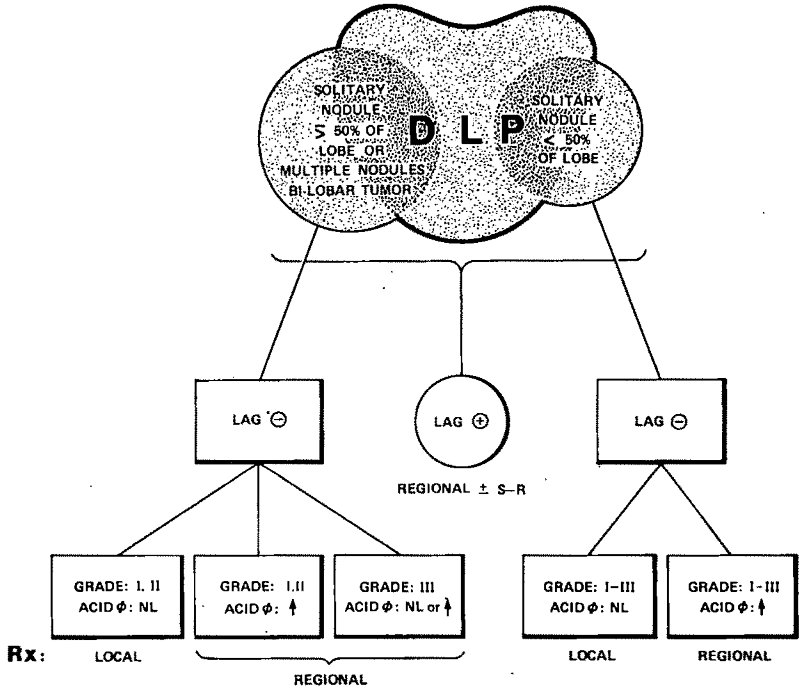


Figure 2 Operational schema depicting suggested treatment approach for patients whose clinical stage is not confirmed by laparotomy and who present with disease limited to the prostate (DLP) as determined by digital rectal examination.

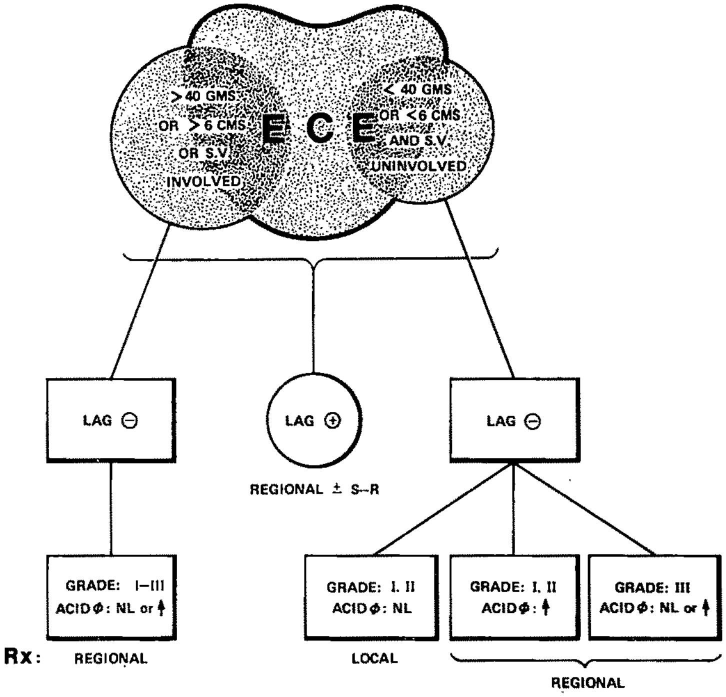


Figure 3 Operational schema for patients with extracapsular extension (ECE) as determined by digital rectal examination.

nodes which are apparently most often involved by tumor (hypogastric and obturator) are not opacified by lymphangiography; (*c*) that false positive lymphangiographic interpretations are relatively rare (49); (*d*) that the incidence of lymph node metastases increases with increasing size of the primary tumor (44, 45), with extracapsular extension (especially seminal vesicle invasion (44, 46), and with histologic dedifferentiation (45); and (*e*) that periaortic node involvement is apparently most always associated with regional lymphadenopathy, but not the converse (49).

In general, the schema reflects our interpretation of the relative value of various clinical parameters in predicting extension of tumor to lymph nodes. We were impressed with the reliability of a positive lymphangiogram. In addition, the presence of extracapsular extension (manifested as obvious palpable tumor or inferred by an elevated acid phosphatase), or histological dedifferentiation appeared to have predictive value and were weighted accordingly. It must be emphasized that the operational schema in many areas represents inferences derived from preliminary data, and will undoubtedly undergo revision as new information becomes available.

It should be noted that according to the schema a significant number of patients with negative lymphangiograms receive regional treatment. The rationale for this reflects the possibility that because of the 20% false negative rate associated with lymphangiography, and the nonopacification of the obturator and hypogastric nodes by this procedure, a negative lymphangiogram does not preclude lymphatic involvement in high risk groups. Therefore, we believe that patients with negative lymphangiograms who have extensive extracapsular disease and/or two or more of the other indicators associated with lymph node metastases should be treated by a regional rather than local approach.

The operational schema is meant to serve only as a general guideline. Actual treatment volume should be adjusted on an individual basis, depending on the patient's general condition, history of prior major abdominal surgery, and/or preexisting GU or GI pathology.

TREATMENT TECHNIQUE (STANFORD)

Extreme care must be taken during the tumor localization procedure for all therapeutic approaches in order to maintain as small a treatment volume as possible if side effects are to be kept at an acceptable level. In general, patients are treated five times per week, receiving 200 rads per fraction.

Local Treatment

Local treatment is delivered using either two 120° right and left lateral arcs or 360° rotational therapy. The former is preferable because of the lateral distribution of the dose, and the latter is reserved for patients with posterior extension of their tumor.

Regional and Supraregional Treatment

Figures 4*A* and *B* show the radiation portal arrangement for the treatment of the prostate, pelvic, and periaortic lymph nodes. A four-field arrangement is used for both the regional and supraregional approach, and all fields are treated daily in order to assure a more homogeneous distribution of dose throughout the treatment volume. In addition, this arrangement decreases the dose to the subcutaneous tissue, thereby avoiding the debilitating subcutanous fibrosis that may occur when coplanar opposed portals or a three-field technique is used. The pelvic field is treated with four parallel opposed portals giving a rectangular dose distribution. The superior level is the superior cortical plate of L_5 (unless a modified regional approach is used), and the inferior extent is at the level of the ischial tuberosities. The periaortic nodes are treated, when indicated, with anterior and posterior opposed and right and left oblique portals. The superior level of the periaortic portal is the superior cortical plate of T_{12}. A calculated gap is introduced between the periaortic and pelvic fields. Patient tolerance permitting, the volume including the periaortic nodes is treated concomitantly with the pelvic treatment volume.

A "split-course" technique is used which introduces a two-week period following an initial dose of 2500 rads to a volume including the prostate and lymph node bearing areas in which only the prostate portal (7 X 7 cm) is treated by rotational therapy. Following this interval an additional 2500 rads is delivered to the original, larger volume. This allows continuous irradiation of the prostate during a seven-week period, while increasing the time to treat the node-bearing areas which are

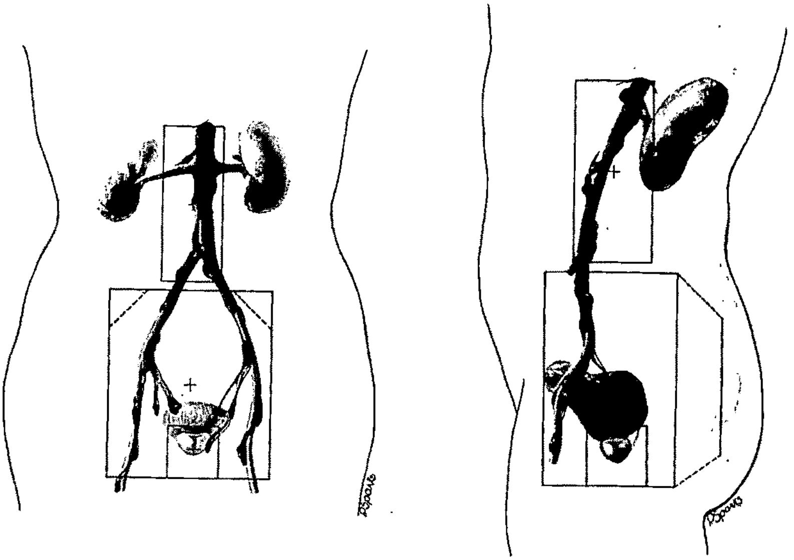


Figure 4 Radiation portals used in the supraregional approach to treat the prostate and the pelvic and periaortic lymph nodes. The point of juncture between the periaortic and pelvic portals is arbitrary. When the regional approach is indicated, the superior level of the pelvic portal coincides with the superior cortical plate of the 5th lumbar vertebral body.

adjacent to the inherently more sensitive intestinal epithelium. Therefore, during a seven-week period, the prostate receives 7000 rads and the nodal volume 5000 rads, including a two-week period in which the majority of the intestine is not irradiated.

FUTURE AREAS OF CLINICAL INVESTIGATION

We believe that although numerous areas of clinical investigation exist, two predominate. The first is the establishment of a randomized, appropriately controlled trial comparing various therapeutic modalities in patients with similar clinical and pathologic stages. It remains to be proven that aggressive treatment, surgery alone, surgery in combination with irradiation, or radiation therapy alone offers better survival rates than conservative treatment in selected groups of patients. The second would test the question whether irradiation to areas of known involvement plus contiguous lymphatics provides better survival rates than irradiation limited to known areas of tumor. The latter question is already being tested by an ongoing randomized trial sponsored by the Divisions of Radiation Therapy and Urology, Stanford University. In addition, this study will provide firm data on the ability of irradiation to sterilize lymphatic metastases. However, since there is no randomization to "prostate-only treatment" in patients with known lymphatic metastases, only inferential data will be obtained as to the ability of irradiation to alter the natural history of prostatic carcinoma once lymphatic metastases have occurred.

Literature Cited

1. Jewett, H. J., Bridge, R. W., Gray, F. G., Shelley, W. M. 1968. The palpable nodule of prostate cancer. Results 15 years after radical excision. *J. Am. Med. Assoc.* 203:403–6
2. Bailar, J. C. III, Byar, D. P., The Veterans Administration Cooperative Urological Research Group. 1970. Estrogen treatment for cancer of the prostate. Early results with 3 doses of diethylstilbestrol and placebo. *Cancer* 26:257–61
3. Budhraja, S. N., Anderson, J. C. 1964. An assessment of the value of radiotherapy in the management of carcinoma of the prostate. *Brit. J. Urol.* 36:535–40
4. Pasteau, O. 1911. Traitement du cancer de la prostate par le radium. *Revue des Maladies de la Nutrition,* pp. 363, 398
5. Young, H. H., Frontz, W. A. 1971. Some new methods in the treatment of carcinoma of the lower genitourinary tract with radium. *J. Urol.* 1:505–34
6. Barringer, B. S. 1917. Radium in the treatment of carcinoma of the bladder and prostate; a review of one year's work. *J. Am. Med. Assoc.* 28:1227–30
7. Deming, C. L. 1922. Cancer of the prostate and seminal vesicles. Results in one hundred cases of cancer of the prostate and seminal vesicles treated with radium. *Surg. Gynecol. Obstet.* 34:99–118
8. Bugbee, H. G. 1921. Experiences with radium in cancer of the prostate. *J. Urol.* 6:459–70
9. Bumpus, H. D. 1922. Roentgen rays and radium in the diagnosis and treatment of carcinoma of the prostate. *Am. J. Roentgenol.* 9:269–89
10. Barringer, B. S. 1942. Prostatic carcinoma. *J. Urol.* 47:306–10
11. Attwater, H. L. 1930. Malignant diseases of prostate and bladder. *Postgrad. Med. J.* 6:6–19
12. Hultberg, S. 1946. Results of treatment with radiotherapy in carcinoma of the prostate. *Acta Radiol.* 27:339–49
13. Smith, G. S., Pierson, E. L. 1930. The value of high voltage X-Ray therapy in carcinoma of the prostate. *J. Urol.* 23:331–41
14. Widmann, B. 1934. Cancer of the prostate. The results of radium and Roentgen-ray treatment. *Radiology* 22:153–59
15. Huggins, C., Stevens, R. E., Hodges, C. V. 1941. C. V. studies on prostatic cancer: Effects of castration on advanced carcinoma of prostate gland. *Arch. Surg.* 43:209–23
16. Flocks, R. H., Kerr, H. D., Elkins, H. B., Culp, D. A. 1952. Treatment of carcinoma of the prostate by interstitial radiation with radioactive gold (^{198}Au): A preliminary report. *J. Urol.* 68:510–22
17. Bagshaw, M. A., Kaplan, H. S. 1962. *Radical External Radiotherapy of Localized Prostatic Carcinoma.* Presented at Tenth Int. Congr. Radiol., Montreal, Canada, Sept. 15–18, 1962
18. Ray, G. R., Cassady, J. R., Bagshaw, M. A. 1973. Definitive radiation therapy of carcinoma of the prostate. A report on 15 years experience. *Radiology* 106:409–18
19. Bagshaw, M. A., Ray, G. R. 1973. Carcinoma of the prostate: The role of radiotherapy. *Proceedings of the XIII International Congress of Radiology, Madrid.* Amsterdam: Excerpta Med.
20. Gardner, A., Bagshaw, M. A., Page, V., Karzmark, C. J. 1972. Tumor localization, dosimetry, simulation and treatment procedures in radiotherapy: The isocentric technique. *Am. J. Roentgenol.* 114:163–71
21. Bagshaw, M. A. 1973. Definitive megavoltage radiation therapy in carcinoma of the prostate. In *Textbook of Radiotherapy,* ed. G. H. Fletcher, 752–67. Philadelphia: Lea & Febiger
22. Cutler, S. J., Ederer, F. 1958. Maximum utilization of the life table method in analyzing survival. *J. Chronic Dis.* 8:699–712
23. Whitmore, W. F. 1963. The rationale and results of ablative surgery for prostatic cancer. *Cancer* 16:1119–31
24. Veterans Administration Co-operative Urological Research Group. 1967. Treatment and survival of patients with carcinoma of the prostate. *Surg. Gynecol. Obstet.* 124:1011–19
25. Bagshaw, M. A., Kaplan, H. S., Sagerman, R. H. 1965. Linear accelerator supervoltage radiotherapy. *Radiology* 85:121–29
26. Del Regato, J. A. 1967. Radiotherapy in the conservative treatment of operable and locally inoperable carcinoma of the prostate. *Radiology* 88:761–66
27. Dykhuizen, R. F., Sargent, C. R., George, F. W., Kunahara, S. S. 1968. The use of cobalt-60 teletherapy in the treatment of prostate carcinoma. *J. Urol.* 100:333–38

28. George, F. W., Carlton, C. E. Jr., Dykhuizen, R. F., Dillon, J. R. 1965. Cobalt-60 telecurietherapy in the definite treatment of carcinoma of the prostate: A preliminary report. *J. Urol.* 93:102–9
29. Bennett, J. E. 1968. Treatment of carcinoma of the prostate by cobalt-beam therapy. *Radiology* 90:532–35
30. Grout, D. C., Grayhack, J. T., Moss, W., Holland, J. M. 1971. Radiation therapy in the treatment of carcinoma of the prostate. *J. Urol.* 100:411–14
31. Odell, R. W., Merill, M. D., Atwood, C. J. 1971. Cobalt-60 teletherapy of localized prostatic carcinoma. A 10 year experience. *J. Urol.* 105:843–46
32. Carlton, C. E. Jr., Dawoud, F., Hudgins, P., Scott, R. Jr. 1972. Irradiation treatment of carcinoma of the prostate: A preliminary report based on 8 years of experience. *J. Urol.* 108:924–27
33. Loh, E. S., Brown, H. E., Beiler, D. D. 1971. Radiotherapy of carcinoma of the prostate: Preliminary report. *J. Urol.* 106:906–9
34. Michaels, M. M., Brown, H. E., Beiler, D. D. 1974. Radiotherapy of carcinoma of the prostate: A followup report. *J. Urol.* 111:72–74
35. Hazra, T. A. 1972. Radiotherapy in the treatment of locally infiltrating carcinoma of the prostate. *J. Urol.* 108:460–61
36. Grossman, I. et al 1974. The early lymphatic spread of manifest prostatic adenocarcinoma. *Am. J. Roentgenol.* 120:673–74
37. Phillips, N. B., Lattimer, J. K. 1972. Complications and regression of primary tumor size with radiotherapy in prostate cancer. *J. Urol.* 108:921–47
38. Rhamy, R. K., Wilson, S. K., Caldwell, W. L. 1972. Biopsy-proved tumor following definite irradiation for resectable carcinoma of the prostate. *J. Urol.* 107:627–30
39. Cox, J. D., Tijerina, A. 1974. Preliminary results of biopsies following irradiation for locally advanced adenocarcinoma of the prostate. *Radiology* 112:215–16
40. Pinck, B. D., Alexander S. 1972. Parathormone potentiated radiophosphorus therapy in prostatic carcinoma. *Urology* 1:201–4
41. Edland, R. W. 1974. Testosterone potentiated radiophosphorus therapy of osseous metastases in prostatic cancer. *Am. J. Roentgenol.* 120:678–82
42. Whitmore, W. F. Jr., Hilaris, B., Grabstald, H. 1972. Retropubic implantation of iodine-125 in the treatment of prostatic cancer. *J. Urol.* 108:918–20
43. Bagshaw, M. A., Ray, G. R., Salzman, J., Meares, E. M. 1975. Extended field radiation therapy of carcinoma of the prostate: A progress report. *Cancer Chemother. Rep.,* Vol. 59, No. 1
44. Flocks, R. H., Culp, D., Porto, R. 1959. Lymphatic spread from prostatic cancer. *J. Urol.* 81:194–96
45. McCullough, D. L., Prout, G. R. Jr., Daly, J. J. 1974. Carcinoma of the prostate and lymphatic metastases. *J. Urol.* 111:65–71
46. Arduino, L. J., Glucksman, M. A. 1962. Lymph node metastases in early carcinoma of the prostate. *J. Urol.* 88:91–93
47. Jewett, H. J. 1970. The case for radical perineal prostatectomy. *J. Urol.* 103:195–99
48. Whitmore, W. F. Jr., MacKenzie, A. R. 1959. Experiences with various operative procedures for the total excision of prostate cancer. *J. Urol.* 12:396–405
49. Castellino, R. A., Ray, G. R., Blank, N., Govan, D., Bagshaw, M. 1973. Lymphangiography in prostatic carcinoma: Preliminary observations. *J. Am. Med. Assoc.* 223:877–81
50. Ray, G. R., Cassady, J. R., Bagshaw, M. A. 1974. External-beam megavoltage radiation therapy in the treatment of post radical prostatectomy, residual or recurrent tumor: Preliminary results. *J. Urol.* In press
51. Gilbertson, V. A. 1971. Cancer of the prostate gland. Results of early diagnoses and therapy undertaken for cure of the disease. *J. Am. Med. Assoc.* 215:81–84
52. Byar, D. P., Mostofi, F. K., The Veterans Administration Cooperative Urological Research Group. 1972. Carcinoma of the prostate: Prognostic evaluation of certain pathologic features in 208 radical prostatectomies examined by the step-section technique. *Cancer* 30:5–13

Copyright 1975. All rights reserved

SUBACUTE SCLEROSING PANENCEPHALITIS

❖7174

Richard S. Johannes, M.D. and John L. Sever, M.D.
Infectious Diseases Branch, National Institute of Neurological Diseases and Stroke,
National Institutes of Health, Bethesda, Maryland 20014

Subacute sclerosing panencephalitis was probably first described by Schilder in 1924 (1) under the name "encephalitis periaxialis diffusa." Other cases were studied by several investigators who, because of the natural variation of the disease and the use of various staining procedures, tended to stress the differences among cases rather than the similarities (2–6). Not until the late 1940s did it become clear that Dawson's inclusion body encephalitis, van Bogaert's subacute sclerosing leucoencephalitis, and the panencephalitis of Petté and Döring were indeed all one and the same disease, now known as subacute sclerosing panencephalitis (SSPE) (7).

CLINICAL SIGNS AND SYMPTOMS

The clinical course of the disease is now usually divided into four stages (8). Stage 1 is often very subtle in its manifestations and includes decreased intellectual skills, rapid swings in mood, inappropriate affect, poor attentiveness in school, and reduced scholastic ability. Drooling and speech changes are occasionally seen in stage 1. Since these changes are usually not indicators of serious organic disease within the age group who contract SSPE, it is not surprising to find that many children are first brought to school principals' offices for discipline and to child psychologists for evaluation before coming to a physician. Stage 2 is heralded by dramatic neurologic signs which belay any question of objective neurologic disease. It is at stage 2 where the physician usually makes first contact with the patient. Myoclonic jerks and stumbling while walking are the hallmarks of the major motor signs. Ataxic, choreoform, and athetotic movements are also seen. Generalized convulsive fits will complicate a small percentage of cases. Ocular signs are seen in approximately 50% of the cases. The signs vary: cortical blindness, optic atrophy, chorioretinitis frequently involving the macula, and extrinsic ocular palsies all have been reported (9). Stage 2 is the level at which the disease shows greatest variability (10). "Spontaneous remission" (11) or perhaps more properly arrest can occur at this time. Although up to 80% of patients in some series (12) perish in 9 months, survival for over 8 years has also been reported (13).

The disease usually passes quickly into stage 3, which is characterized by marked mental loss, loss of speech function, coma, opisthotonus, extensor plantar responses, decerebrate rigidity, and decorticate posturing. The coma may last several months. Derangements in autonomic nervous function such as pallor, flushing, and instability of temperature homeostasis are also seen. If the patient does not succumb to the complications of long term coma, usually respiratory or urinary infection, he passes into stage 4. The stage 4 patients may seem improved, but in fact a state of near total loss of cortical function exists. Startle responses to noise may be seen but other than this, very little integrated nervous system function remains. Wandering eye movements, episodic laughing or crying, hypotonia, and flexing of the arms and legs with turning of the head to one side complete the picture of stage 4 SSPE. The derangements of autonomic function and complications of care for such a neurologically crippled patient usually lead to the patient's demise, and death ensues due to vasomotor collapse and/or infectious complications.

EPIDEMIOLOGY

Due to the relative infrequency of SSPE and the possible decline in incidence, current epidemiologic data are scant. Though exact numbers vary, there appears to be a male preponderance with attack rates of up to 3.3:1 (8). The geographic data suggest that there is a proclivity to occur in the southeastern United States. It occurs with an incidence of approximately one per million in the United States (14). More than 50% of the patients have a history of clinical measles at less than two years of age (14), suggesting that immaturity of the immune system may be a contributing factor (15) (see Figure 1). The clinical onset usually occurs between ages five and twelve with a mean of near ten (see Figure 2). The "incubation period" following clinical measles is usually six years (16). There is evidence (14, 17) in favor of an increased incidence in rural areas. The incidence in cities with populations greater than one million is 0.475 per million, but this is confounded by the fact that there are few cities with populations in excess of one million in areas where SSPE is most prevalent (17).

DIAGNOSIS

Prior to the demonstration of recoverable measles virus from brain tissue of patients with SSPE, much circumstantial evidence was accrued to implicate the measles virus as the etiologic agent (1). It was from this work that practical tests for diagnosis were developed (18). The most useful and specific clinical tests involve serologic and biochemical examination of cerebrospinal fluid (CSF) and serum. Routine examination of CSF from patients with SSPE rapidly disclosed a high frequency of patients with an elevated gamma globulin level and paretic colloidal gold curve (19, 20). The globulin level is often elevated well beyond the normal range of less than 10% of the total protein (21). Finding a normal level does not exclude

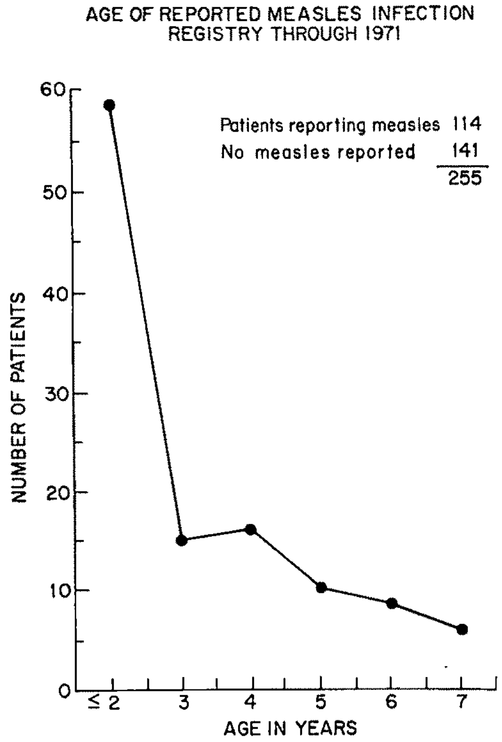


Figure 1 Age of reported measles infection registry through 1971. Figure courtesy of Dr. J. T. Jabbour.

the diagnosis. Using albumin as a marker of the integrity of the blood-brain barrier, Tourtellotte (20) was able to calculate that 95% of the increased gamma globulin was due to local production within the central nervous system (CNS). Other investigators (22), using intravenous administration of radioisotopically labeled immunoglobulin G, demonstrated by kinetic analysis that the major portion of CSF immunoglobulin G in patients with SSPE was derived from a nonvascular source. This added credence to the concept of endogenous synthesis of immunoglobulin G within the CNS of SSPE patients. Electrophoretic separation of SSPE spinal fluid frequently shows a pattern of from one to five distinct bands migrating to the far cathodic region of the globulin smear (23). This pattern is called the oligoclonal aspect and is seen in few other diseases, e.g. CNS syphilis, multiple sclerosis, and CNS trypanosomiasis. There is increasing evidence that these bands represent homogeneous groups of measles-specific antibody (24, 25). It was ultrastructural examination suggesting myxo- or paramyxovirus viral inclusions coupled with elevated CSF gamma globulin levels that led to investigations of viral antibody titers in patients with SSPE. Connolly described three patients in 1967 with elevated titers

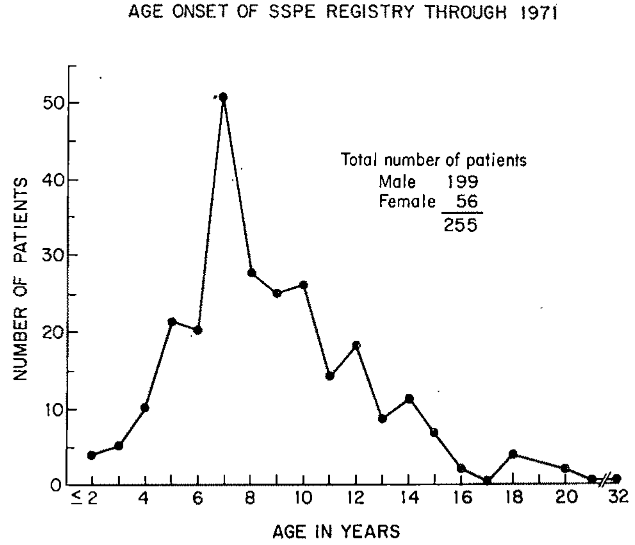


Figure 2 Age onset of SSPE registry through 1971. Figure courtesy of Dr. J. T. Jabbour.

to measles in both serum and CSF (26, 27). Connolly's observation was rapidly supported by other investigators (28–33). The high serum titer to measles in patients with SSPE was surprising, often being higher than the titers observed in recovery from acute natural measles (28, 33). The specificity of elevated measles titers was strengthened by the demonstration that the ratios of CSF to serum titers of measles were significantly reduced compared to polio titers (28), and by the demonstration that only the measles titer remained elevated at high titer on longitudinal examination of two patients with 13 different viral antigens over a seven-month period (32). The corollary to the experiments using complement fixation (CF) and hemagglutination inhibition (HI) titers was the demonstration of measles antigen within the inclusion bodies of neurons and oligodendrocytes by immunofluorescence (34). Though immunofluorescence is very sensitive, it does require brain biopsy. Determination of serum and CSF measles titers requires only a venipuncture and a lumbar puncture. Due to the relative ease and great specificity of these simpler tests, they have virtually replaced the need for brain biopsy in making the diagnosis (36). Of the four tests, HI and CF done on both serum and CSF, the CF test on CSF is clearly the best single test. However, reliance on a single test will miss a percentage of cases (35). Currently, only six states provide all four tests. Twenty two states provide CF testing CSF. A complete listing of the tests available in the various states has recently been reviewed (35).

The electroencephalogram (EEG) can be very helpful in making the diagnosis. Examination of the EEG also has helped to reveal that the measles virus may exhibit more neurotropism than was once appreciated. In one study of 680 patients with clinical measles without any outward evidence of CNS involvement, 51% of patients demonstrated abnormalities in their EEGs during the acute or postacute phase of their illness (37). In the same study, 37 patients from the total population of 717 had clinical encephalitis. Multinucleated giant cells and typical inclusions were seen in pathologic specimens. In addition, the measles virus has been isolated from brain specimens of patients with measles encephalitis (38), suggesting that this illness may represent true measles infection of the brain and not simply an immunologically mediated response to measles virus.

In SSPE, the EEG pattern of suppression, burst is characteristic. The pattern consists of high amplitude slow delta waves occurring at regular 3-sec intervals (8) (see Figure 3). These findings are, however, variable and repeat tracings may be required to demonstrate the typical EEG pattern. Once seen, the suppression burst pattern is quite specific. The EEG may revert to normal after abnormal records have been obtained (11). In one series (39), 6 of 11 patients examined an average of six months after onset of disease and normal EEGs. One patient had a normal EEG at the time of well developed myoclonus. Though all patients eventually developed abnormal encephalograms, only 4 showed the classical suppression burst pattern.

PATHOLOGY

It was from close inspection of the cellular changes in pathologic brain specimens obtained both by biopsy and necropsy that the clues to viral pathogenesis of SSPE were manifest. Among the early investigations of SSPE brain, it was Dawson in 1933 and 1934 (2, 3) who called attention to eosinophilic Cowdry type-A inclusion bodies in the cytosol and nucleus of cortical neurons. Both oligodendroglia and neurons are involved in the process. With the exception of inclusion bodies the pathologic picture is not unlike that of any primary viral encephalitis.

It has been noted that there is more white matter demyelination and sclerosis in late cases while intranuclear inclusions are more readily demonstrable in cases of short duration (40). The histology is one of varied neuronal loss with occasional neuronophagia, leptomeningeal infiltration with lymphyocytes and plasma cells, perivascular cuffing in both central white matter and throughout the cortex, and scattered microglial nodules. If examined post mortem, the inflammatory reaction may be minimal or nonexistent with neurons showing Alzheimer's neurofibrillary changes (40). The earliest change is the appearance of small intranuclear inclusion bodies approximately 1–2 μm in diameter (41). They are frequently multiple and differentiable from the nucleolus by their eosinophilic straining characteristics. These bodies do not stain with the Feulgen reaction for DNA but rather show intense pyroninophilic staining, indicating a high content of RNA. The inclusions occasionally measure up to 3 μm and are no longer differentiable from Cowdry type-A inclusions. On ultrastructural analysis these inclusions termed the nucleoliform inclusions appear as collections of loosely packed granules measuring 300–500 Å. The Cowdry type-A particles enlarge, nearly filling the nucleus and leaving a pale halo surrounding the inclusion. These inclusions are a second type called the multitubular inclusions due to their appearance on electron microscopy

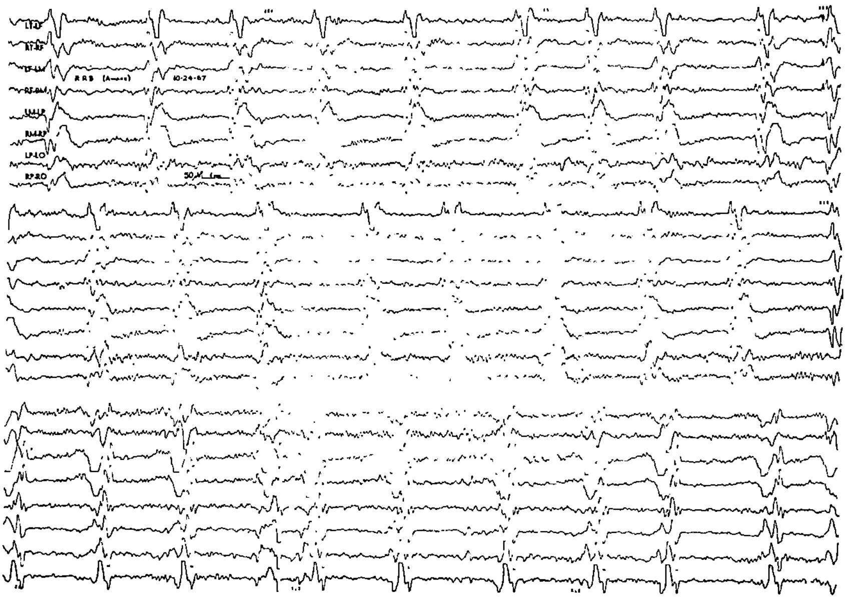


Figure 3 Typical electroencephalogram in subacute sclerosing panencephalitis. Periodic bursts of high voltage activity are demonstrated. Reproduced from *J. Am. Med. Assoc.* 207:2248 (1969).

(40). The multitubular inclusions show much less staining with pyronin but have a high affinity for protein stains. Multitubular inclusions are collections of tubules with an outside diameter of 170–190 Å and an inside diameter of 40–60 Å. They measure up to 12 μm. The tubules are usually randomly dispersed but occasionally are closely packed as a pseudocrystalline array. The third type of inclusion, the cytoplasmic inclusions, vary in size but ultrastructurally are very similar to the multitubular inclusions seen in the nucleus. They consist of the 180-Å tubules with a superimposed course and dense granular component. The cytoplasmic inclusions show high affinity for both pyronin and bromosulfalein, indicating the presence of both RNA and protein (40, 42–44). By using immunochemistry several investigators have shown that immunoglobulin G, immunoglobulin M, and complement (45–47) are bound by the cellular inclusions in SSPE brain tissue. Neutralizing antibody against measles virus has been eluted in low concentration from SSPE tissue and was shown to be directed against the nucleoprotein core of the virus and not against any of the surface antigens (47).

ETIOLOGY

It was the clues from electron microscopy and serology (discussed under diagnosis) which led investigators to make attempts to isolate a myxovirus or paramyxovirus particle from patients with SSPE. Many techniques were tried by investigators only to end in failure (30, 48–50). Eventually measles antigen, syncytia formation, and recovery of fully infectious virus was achieved using the technique of co-culture of SSPE brain tissue with HeLa cells (51–53).

Thus, it became clear that SSPE is a slow virus infection of the brain caused by measles virus (54). What is not as yet clear is what events allow the rare patient who is exposed to wild measles to develop a smoldering, progressive infection within the CNS. Central to this issue is whether the virus is different from the usual measles virus or whether the host's response to a common agent is somehow varied from normal so as to allow a chronic infection to occur.

Genetic susceptibility is unlikely, as the disease has occurred in only one of a pair of identical twins (55). The age of measles infection (14, 15) and/or simultaneous infection with other viruses such as varicella (17) seem to be important factors in the pathogenesis of SSPE.

SSPE may be more of a systemic disease than is clinically apparent. With the use of indirect immunofluorescence, measles antigen has been found in lung, kidney, and spleen in addition to brain (47). Co-cultivation techniques have rescued live measles virus from lymph nodes of patients with SSPE in some cases (56). To date there are no viral isolations from the thymus glands of SSPE patients (57).

The virus recovered from patients with SSPE is a measles-like virus which shows an antigenic composition which differs from both the wild-type measles and the Edmonston vaccine strain (58). As more strains of SSPE virus are being isolated from patients it is becoming apparent that there is significant biologic and antigenic heterogeneity of the viral SSPE isolates (58, 59), as is the case with both wild and vaccine strains of measles. For example, Hamilton & Barbosa reported on a comparison of two strains of SSPE and Edmonston strain (60). They showed that one of the SSPE strains (Dean) was similar to Edmonston when examined by growth characteristics, morphology, infectivity, and antigenicity. The other SSPE (Halle) strain was markedly different from both the Edmonston strain and the DEAN strain. There is little doubt that significant biologic changes in SSPE virus occur as a consequence of sequential in vivo passage while infecting human brain, not to mention the possible effects of cellular fusion during co-culture required to rescue live virus. When antibody assays from patients with SSPE were measured by complement fixation of hemagglutination inhibition, similar titers were obtained whether SSPE or Edmonston strain antigens were used in the assay (58, 59). However, the neutralization titers revealed a 4- to 8-fold decrease in titer of SSPE sera assayed against SSPE virus as opposed to wild-type measles virus (59). This is consistent with the findings of greater CF titers than HI titers in patients with SSPE, and has been interpreted to represent antibody response elicited against the nucleoprotein core as opposed to the site of neutralizing antibody, namely the outer envelope (35).

One very interesting strain has been isolated in Japan by Doi et al (61). This particular variant of SSPE virus appears to be absolutely cell associated. It causes isolated plaques in tissue culture as virus seemingly spreads from cell to cell without cycles of liberation of free virus into the supernatant fluid with subsequent infection of nearby cells. Ultrastructurally the virus causes several differences compared with a productive SSPE measles infection in the same cell line (62). However, the virus does induce surface membrane changes. The membrane changes with the Japanese strain appear continuous along the cell surface, and not as discontinuous areas with alignment of tubules followed by formation of mature budding particles as seen in productive infection. Though no free virus is formed, anti-measles serum will block infection of tissue culture if applied prior to cellular fusion. Hence, it seems that the measles virus is expressed in an incomplete form at the level of the plasmalemma.

One other question which has arisen with respect to SSPE virus is whether it is a paramyxovirus closely related to measles, but which resides primarily in an animal reservoir. This search was promulgated by the relative rural distribution of SSPE cases. Indeed antibody to canine distemper virus is present in patients with SSPE (63). Though such antibody is seen in patients convalescing from wild measles infection, the titer is significantly increased in patients with SSPE (63). The relevance of this is unclear as distemper virus causes a homotypic antibody response whereas measles infection frequently leads to positive, albeit low, titers to distemper (32). In addition, superinfection with canine distemper is a possibility (63). There is evidence for such a suggestion since the Japanese strain of Doi did not block superinfection with other viruses in chronic in vitro SSPE viral infection (61). Transfer of vaccine measles, SSPE virus, and canine distemper virus to lambs and calves is relevant to this problem. Attempts to do this resulted (64) in neurologic symptoms in a few of those animals given SSPE virus. The animals given vaccine measles and canine distemper developed homotypic antibody but no neurologic symptoms. On the other hand, the animals given SSPE virus did not show any HI antibody response, but measles antigen could be demonstrated in brain material by fluorescent antibody techniques. Whether infection from an animal reservoir carrying a variant of measles is relevant to human SSPE is an unresolved point at present and further study will be needed to answer the question.

The nature of the host immune response in SSPE continues to be a source of confusion. Earlier it was felt that cellular immunologic abnormalities were present in patients with SSPE. Gerson & Haslam (65) reported depressed cellular immune function to six common skin test antigens, dinitrochlorobenzene responsiveness, and delayed allogenic skin rejection in four patients with SSPE. They also showed decreased germinal center activity in lymph nodes following diphtheria-tetanus toxoid boosters. The question arose as to the specificity of these immunologic anomalies and indeed whether they were primary or secondary events in pathogenesis. Good et al (66) reported on several children with congenital agammaglobulinemia who were unable to produce antibodies and had only insignificant levels of immunoglobulins in the circulation. On contracting measles, these children showed a normal progression of signs and symptoms and were subsequently immune to reinfection, suggesting an important role for cellular immunity in measles infections. Enders (67) reported on leukemic children who were immunosuppressed with cortisone as part of their treatment, and in whom measles may take the form of a virulent giant-cell pneumonia. Furthermore, Hanissian et al (68) reported a case of subacute encephalitis clinically resembling SSPE in a patient with X-linked hypogammaglobulinemia. The boy's illness was consistent with SSPE by clinical and EEG findings. However, he lacked any measles antibody in both serum and CSF by both CF and HI techniques. Although no virus was recovered from the brain, electron microscopy showed occasional cells with nucleoliform inclusions and fluorescent microscopy showed measles antigen in approximately 5% of primary brain culture cells. This report supports the tenet that cellular immune mechanisms play a significant role in the development of SSPE.

It was reported initially that SSPE patients' leukocytes did not show migration inhibition or blastic response when exposed to cell-free measles antigen, suggesting lymphocyte unresponsiveness to measles virus (69). However, there have since been reports (70–72) which offer strong evidence in favor of the immune competence of lymphocytes with respect to measles in patients with SSPE. The requirement for demonstrating a measles-specific immune lymphocyte response was the use of measles infected cells and not simply free measles virus (70). Ahmed et al (71) have recently reported on a high molecular weight heat-sensitive sialo-protein or protein complex which was able to block the expression of measles-sensitive lymphocytes from SSPE patients. The inhibitory substance was present in the cerebrospinal fluid at ten times serum concentration although the absolute amount of protein is nearly two orders of magnitude lower within the CSF. They were also able to demonstrate that the inhibitory substance was measles-specific and did not block elaboration of migration inhibitory factors when tested against several other antigens.

ANIMAL MODELS

Although SSPE appears to be a uniquely human disease, animal models of SSPE are available. In one model (73, 74), suckling hamsters from measles-immunized mothers were given Schwartz strain measles virus intracerebrally. These baby hamsters had high levels of maternal antibody acquired transplacentally. Of these animals, 83% did not develop acute encephalitis as did 100% of baby hamsters from nonimmune mothers. A small number of the passively immunized animals exhibited neurologic symptoms, and these animals had pathologic and tissue culture evidence of measles infection. In addition, when normally appearing, symptom-free hamsters which had survived 51 days postinoculation were given cyclophosphamide, a potent immunosuppressive agent, one third of them developed an acute measles encephalitis. Measles virus could be recovered from the brains of these animals as long as 90 days postinoculation. Another investigation (15) involved intracerebral inoculation of an SSPE strain into hamsters of varied ages. Adult animals developed an acute, nonfatal, self-limited inclusion-cell encephalitis. Newborn animals developed a uniformly fatal giant-cell encephalitis. However, weanling hamsters of 21 to 22 days of age showed a varied response. Some developed a subacute illness while others developed a chronic illness with survival throughout the entire 120-day course of the study. Cell-free virus was demonstrated in animals sacrificed up to 8 days postinoculation. However, cell-free virus could not be shown in those animals who developed chronic illness with neurologic signs, and co-cultivation techniques were required to rescue cell-associated virus from these animals (15, 64). Although the animal models do not provide perfect corollaries to human SSPE, these models may provide great insight into the mechanisms of persistent measles infection.

TREATMENT

Several efforts in treating SSPE with amantadine, antimetabolites, and steroids (9, 75–77) have been made but none has shown any real promise. Two approaches currently under active investigation may change this. One involves the use of Lawrence's transfer factor, which is able to correct cellular immune deficits in some models (78). However, the best to be expected from this would be arrest of the disease at the stage of institution of therapy. Furthermore, if indeed the lymphocytes of patients suffering from SSPE are competent and their expression is being blocked by a circulating inhibitory factor as suggested by Ahmed et al (70), no response to transfer factor would be expected. Secondly, there is a single case report of improvement following institution of treatment with the agent isoprinosine (79). Whether the patient has truly demonstrated a response to treatment or is in remission as has been the case with other long term improvements (80) will require further time and investigation. At the time of this publication the patient has remained stable and is continuing to go to school (R. H. Mattson, personal communication).

FUTURE

Looking ahead into what may come of investigation of SSPE in the next few years, several interesting questions should be solved. First of all, can the disease be eliminated? It seems clear, at present, that no dramatic increase of SSPE has occurred as a consequence of mass immunization with the live vaccine. In fact, the incidence of SSPE may indeed be declining in the United States. Thus, continued mass vaccination may eliminate the disease. Seven years have now passed since the introduction of mass immunization against measles. Since the incubation period for SSPE is six years, any effects of mass immunization on the incidence of SSPE should soon become evident. Second, the role and mechanism of latency of paramyxoviruses may be elucidated more precisely as a result of further examination of the curious Japanese strain as well as by continued research into existing animal models for SSPE. Perhaps the most exciting advances may come from the elucidation of the nature and mechanism of the recently described measles-specific inhibitory substance which appears able to block expression of cellular immunity in SSPE.

Literature Cited

1. Schilder, P. 1924. Die Encephalitis periaxialis diffusa (nebst Bemerkungen über die Apraxie des Lidschlusses. *Arch. Psychiat. Nervenkr.* 71:327
2. Dawson, J. R. Jr. 1933. Cellular inclusions in cerebral lesions of lethargic encephalitis. *Am. J. Pathol.* 9:7
3. Dawson, J. R. Jr. 1934. Cellular inclusions in cerebral lesions of epidemic encephalitis. *Arch. Neurol. Psychiat.* 31:685
4. Petté, H., Döring, G. 1939. Über einheimische panecephalomyelits vom charakter der encephalitis japonica. *Deut. Z. Nervenheilk.* 149:7
5. Van Bogaert, L., De Busscher, J. 1939. Sur la sclèrose inflammatoire de la substance blanche des hemispheres (Spielmeyer). *Rev. Neurol.* 71:679
6. Van Bogaert, L. 1945. Une leucoencéphalite sclèrosante subaique. *J. Neurol. Neurosurg. Psychiat.* 8:101
7. Adams, J. M. 1970. Persistence of measles virus and demyelinating disease. *Hosp. Pract.* 1:87
8. Jabbour, J. T. et al 1969. SSPE: A multidisciplinary study of eight cases. *J. Am. Med. Assoc.* 27:2248
9. Hiatt, R. L., Grizzard, H. T., McNeer, P., Jabbour, J. T. 1971. Opthalmologic manifestations of SSPE. *Trans. Am. Acad. Ophthalmol. Otolaryngol.* 74:344
10. Lowenthal, A., Moya, G., Poire, R., Macken, J., DeSmedt, R. 1972. SSPE: A clinical and biological reappraisal. *J. Neurol. Sci.* 15:267
11. Donner, M., Waltimo, O., Porras, J., Forsius, H., Saukkonen, A. L. 1973. Subacute sclerosing panencephalitis as a cause of chronic dementia and relapsing brain disorder. *J. Neurol. Neurosurg. Psychiat.* 35:180
12. Chao, D. 1962. Subacute inclusion body encephalitis. *J. Pediat.* 61:501
13. Landau, W. M., Luse, S. A. 1958. Relapsing inclusion encephalitis (Dawson type) of eight years duration. *Neurology* 8:669
14. Jabbour, J. T., Duenas, D. A., Sever, J. L., Krebs, H., Horta-Barbosa, L. 1972. Epidemiology of SSPE. *J. Am. Med. Assoc.* 220:959
15. Byington, D. P., Johnson, K. P. 1972. Experimental SSPE in the hamster: Correlation of age with chronic inclusion cell encephalitis. *J. Infec. Dis.* 126:18
16. Sever, J. L., Ellenberg, J. H., Krebs, H. M., Jabbour, J. T. 1974. Subacute sclerosing panencephalitis and measles vaccine. *Pediat. Res.* 8:418
17. Detels, R., Brody, J., McNew, J., Edgar, A. 1973. Further epidemiological studies of SSPE. *Lancet* 2:11
18. Dietzman, D. et al 1972. Diagnosis of subacute sclerosing panencephalitis by a simple spinal fluid gel precipitation test for measles. *Pediatrics* 49:133
19. Kolar, O. 1968. Immunopathologic aspects of subacute sclerosing panencephalitis. *Neurology* 18:107
20. Tourtellotte, W. W., Parker, J. A., Herdon, R. M., Cuadros, C. V. 1968. Subacute sclerosing Panencephalitis: Brain immunoglobulin-G, measles antibody and albumin. *Neurology* 18:117
21. Harrison, T. R. 1974. *Harrison's Textbook of Medicine,* p. 2039. New York:McGraw-Hill. 7th ed.
22. Cutler, R. W. P., Merler, E., Hammerstad, J. P. 1968. Production of antibody by the central nervous system in subacute sclerosing panencephalitis. *Neurology* 18:129
23. Connolly, J. 1972. Subacute sclerosing panencephalitis. *J. Clin. Pathol.* 6:73
24. Link, H., Norrby, E., Olsson, G. E. 1973. Immunoglobulins and measles antibody in optic neuritis. *N. Engl. J. Med.* 289:1103
25. Norrby, E., Link, H., Olsson, J. E. 1974. Measles virus antibodies in multiple sclerosis. *Arch. Neurol.* 30:285
26. Connolly, J. H., Allen, I. V., Hurwitz, L. J., Millar, H. D. 1967. Measles virus antibody and antigen in subacute sclerosing panencephalitis. *Lancet* 1:542
27. Connolly, J. H., Haire, M., Hadden, D. S. M. 1971. Measles immunoglobulins in subacute sclerosing panencephalitis. *Brit. Med. J.* 1:22
28. Connolly, J. H. 1968. Additional data on measles virus antibody and antigen in subacute sclerosing panencephalitis. *Neurology* 18:87
29. Griffith, J. F., Katz, S. L. 1968. Subacute sclerosing panencephalitis, laboratory findings in 6 cases. *Neurology* 18:95
30. Adels, B. R., Gajdusek, D. C., Gibbs, C. J. Jr., Albrecht, P., Rogers, N. 1968. Attempts to transmit subacute sclerosing panencephalitis and isolate a measles related agent, with a study of the immune response in patients and experimental animals. *Neurology* 18:30

31. Kennedy, C. 1968. A ten-year experience with subacute sclerosing panencephalitis. *Neurology* 18:58
32. Sever, J. L., Zeman, W. 1968. Serological studies of measles and subacute sclerosing panencephalitis. *Neurology* 18:95
33. Berman, P. H., Giles, J. P., Krugman, S. 1968. Correlation of measles and subacute sclerosing panencephalitis. *Neurology* 18:91
34. Lennette, E. H., Magoffin, R. L., Freeman, J. M. 1968. Immunologic evidence of measles virus as an etiologic agent in subacute sclerosing panencephalitis. *Neurology* 18:21
35. Furesz, J. 1971. Serologic diagnosis of measles in subacute sclerosing panencephalitis. *N. Engl. J. Med.* 284:729
36. Sever, J. L., Krebs, H., Ley, A., Horta-Barbosa, L., Rubenstein, D. 1974. Diagnosis of subacute sclerosing panencephalitis: The value of and availability of measles antibody determination. *J. Am. Med. Assoc.* 228:604
37. Gibbs, F. A., Carpenter, P. R., Spies, H. W. 1969. Electroencephalographic abnormality in "uncomplicated" childhood measles. *J. Am. Med. Assoc.* 171:1050
38. Ter Meulen, V., Muller, D., Käckell, M., Katz; M., Meyermann, R. 1972. Isolation of infectious measles virus in measles encephalitis. *Lancet* 2:1172
39. Sisk, M., Griffith, J. 1974. Normal electroencephalograms in subacute sclerosing panencephalitis. *Pediat. Res.* 8:429
40. Herndon, R. M., Rubinstein, L. J. 1968. Light and electron microscopy observations on the development of viral particles in the inclusions of Dawson's encephalitis (subacute sclerosing panencephalitis). *Neurology* 18:8
41. Bootelle, M., Fonatine, C., Vedrenne, C., Delarve, J. 1965. Sur un cas d'encèphalite subaiqüe à inclusions. Etude anatomo-clinique et ultra structurale. *Rev. Neurol.* 118:454
42. Shinsaku, O., Ter Meulen, V., Muller, D., Katz, M., Koprowski, H. 1970. Electron microscopic observations in SSPE brain cell cultures. Their correlation with cytochemical and immunocytological findings. *J. Virol.* 3:370
43. Shaw, O. 1968. Electron microscopic observations in subacute sclerosing panencephalitis: A supplementary report. *Neurology* 18:138
44. Zu Rhein, G. M., Chou, S. M. 1968. Subacute sclerosing panencephalitis. Ultrastructural study of a brain biopsy. *Neurology* 18:146
45. Ter Meulen, V., Müller, D., Joppich, G. 1969. Immunohistological, microscopical and neurochemical studies on encephalitides. *Acta Neuropathol.* 12:244
46. Sharp, G. 1971. Autoantibodies and compliment in SLE: A re-examination. *Hosp. Pract.* 1:109
47. Jenis, E. N., Kniesser, M. K., Rothouse, P., Jensen, G., Scott, R. 1973. SSPE: Immunoultrastructural localization of measles-virus antigen. *Arch. Pathol.* 95:81
48. Harter, D. H., Tellez-Nagel, I. 1968. Attempts to isolate SSPE agent in cell culture. *Neurology* 18:133
49. Perier, O., Thirty, L., Vanderhaeghen, J. J., Pelc, S. 1968. Attempts at experimental transmission and electron microscopic observations in subacute sclerosing panencephalitis. *Neurology* 18:138
50. Burnstein, T., Byington, D. P. 1968. On the isolation of measles virus from infected brain tissue. *Neurology* 18:162
51. Horta-Barbosa, L., Fuccillo, D., Sever, J. L. 1969. SSPE: Isolation of measles virus from a brain biopsy. *Nature* 221:974
52. Horta-Barbosa, L. et al 1969. Isolation of measles virus from brain cell cultures of two patients with SSPE. *Proc. Soc. Exp. Biol. Med.* 132:272
53. Payne, F., Baublis, J., Stabashe, H. 1969. Isolation of measles virus from cell cultures of brain tissues from a patient with SSPE. *N. Engl. J. Med.* 281:585
54. Measles—quick and slow. *Lancet* 2:27, 1971
55. Whitaker, J., Sever, J., Engel, W. K. 1972. SSPE in only one of identical twins. *N. Engl. J. Med.* 287:864
56. Horta-Barbosa, L., Hamilton, R., Wittig, B., Fuccillo, D., Sever, J. L. 1971. Subacute sclerosing panencephalitis: Isolation of suppressed measles virus from lymph node biopsies. *Science* 173:840
57. Burnet, F. M. 1968. Measles as an index of immunologic function. *Lancet* 1:610
58. Horta-Barbosa, L. et al 1970. Some characteristics of SSPE measles virus. *Proc. Soc. Exp. Biol. Med.* 134:17
59. Payne, F., Baublis, J. 1973. Decreased reactivity of SSPE strains of measles virus with antibody. *J. Infec. Dis.* 127:505
60. Hamilton, R., Barbosa, L., Dubois, M.

1973. SSPE virus study of biological markers. *J. Virol.* 12:632
61. Doi, Y. et al 1972. Properties of a cytopathic agent isolated from a patient with SSPE in Japan. *Jap. J. Med. Sci. Biol.* 25:321
62. Dubois-Dalcq, M., Horta-Barbosa, L., Hamilton, R., Sever, J. L. 1974. Comparison between productive and latent SSPE viral infection in vitro. *Lab. Invest.* 30:241
63. Sato, T. A., Yamanouchi, K., Shishido, A. 1973. Presence of neutralizing antibody to canine distemper virus in sera of patients with SSPE. *Arch. Gesamte Virusforsch.* 42:36
64. Thein, P. et al 1972. SSPE: Transmission of the virus to calves and lambs. *Arch. Neurol.* 27:540
65. Gerson, K., Haslam, R. 1971. Subtle immunologic abnormalities in four boys with SSPE. *N. Engl. J. Med.* 285:78
66. Good, R. A., Zak, S. J. 1956. Disturbances in gamma globulin synthesis as "Experiments of nature." *Pediatrics* 18:109
67. Enders, J. F., McCarthy, K., Mitus, A., Cheatham, W. J. 1959. Isolation of measles virus at autopsy in cases of giant-cell pneumonia without rash. *N. Engl. J. Med.* 261:875
68. Hanissian, A., Jabbour, J. T., DeSamerens, S., Garcia, J., Horta-Barbosa, L. 1972. Subacute encephalitis and hypogammaglobulinemia. *Am. J. Dis. Child.* 23:151
69. Moulias, R. S., Reinert, P., Goust, J. M. 1971. Immunologic abnormalities in SSPE. *N. Engl. J. Med.* 285:1099
70. Thurman, G. B. et al 1973. Lymphocyte activation in SSPE virus and CMV infections. *J. Exp. Med.* 138:839
71. Ahmed, A. et al 1974. Demonstration of a blocking factor in plasma and spinal fluid of patients with SSPE. *J. Exp. Med.* 193:902
72. Mizutani, H., Saito, S., Niehi, K., Izuchi, T. 1971. Cellular hypersensitivity in SSPE. *J. Am. Med. Assoc.* 216:1201
73. Wear, D. J., Rapp, F. 1971. Latent measles virus infection of the hamster CNS. *J. Immunol.* 107:1593
74. Knight, P., Duff, R., Rapp, F. 1972. Latency of human measles virus in hamster cells. *J. Virol.* 10:995
75. Haslam, R., McQuillen, M., Clark, D. 1969. Amantadine therapy in SSPE. *Neurology* 19:1080
76. Joncas, J. et al 1973. SSPE: Elevated Barr-Epstein virus antibody titers and failure of amantadine therapy. *J. Neurol. Sci.* 21:381
77. Freeman, J. M. 1968. Treatment of subacute sclerosing panencephalitis with 5-bromo-2-deoxyuridine and pyran copolymer. *Neurology* 18:176
78. Lawrence, H. S. 1971. Transfer factor and cellular immunity. *Immunobiology,* Chap. II, 104. Stamford, Conn.: Sinauer
79. Mattson, R. H. 1974. Subacute sclerosing panencephalitis recovery associated with isoprinosine therapy: Report of a case. *Neurology* 24:383
80. Resnick, J. S., Engel, W. K., Sever, J. L. 1968. Subacute sclerosing panencephalitis, spontaneous improvement in a patient with measles antibody in blood and spinal fluid. *N. Engl. J. Med.* 279:126

CUMULATIVE INDEXES

CONTRIBUTING AUTHORS VOLUMES 22-26

A

Abelmann, W. H., 24:145-52
Aberdeen, E., 26:451-64
Arnone, A., 25:123-30
Atkinson, A. J., 25:99-113

B

Bagshaw, M. A., 26:567-88
Baldessarini, R. J., 23:343-54
Barchas, J. D., 26:499-520
Barker, W. H. Jr., 25:75-81
Barlow, G. H., 24:335-44
Barnes, A. C., 22:133-44
Bausher, J. C., 23:39-56
Becker, M. A., 25:15-28
Beisel, W. R., 26:9-20
Bellingham, A. J., 22:221-34
Berman, M. L., 26:391-401
Bernard, J., 25:39-51
Biggar, W. D., 24:135-44
Bitensky, M. W., 23:263-84
Blum, C. B., 26:365-71
Blumberg, B. S., 23:161-76
Boshes, B., 26:465-70
Bouhuys, A., 24:111-34
Braverman, L. E., 26:443-49
Breslow, J. L., 24:315-24
Brewer, G. J., 25:29-38
Brummer, D. L., 25:115-22
Brunner, E. A., 26:391-401
Buckner, C. D., 25:179-98
Burch, G. E., 22:327-42
Burchenal, J. H., 23:77-92
Burke, J. F., 24:289-94

C

Cadoret, R. J., 26:21-25
Canterbury, J. M., 24:217-32
Chase, T. N., 26:419-26
Cheng, S. C., 26:391-401
Childs, B., 23:373-406
Coburn, R. F., 24:241-50
Cohen, S., 26:373-90
Collen, D., 25:447-55
Collinsworth, K. A., 25:143-48
Coombs, J., 23:149-60
Cooperband, S. R., 22:185-220
Cotzias, G. C., 22:305-26
Creger, W. P., 24:75-82
Cuatrecasas, P., 24:233-40

D

Dale, D. C., 26:277-81
Dao, T. L., 23:1-8
Davison, R., 25:99-113
Deftos, L. J., 25:323-31
DeGraff, A. C. Jr., 24:111-34
DeLuca, H. F., 25:349-67
DePasquale, N. P., 22:327-42
Desbuquois, B., 24:233-40
DeVries, W. C., 26:203-17
Dixon, F. J., 25:83-98
Doherty, J. E., 26:159-71
Dowling, M. D. Jr., 23:77-92
Downing, D. T., 26:27-32
Dreizen, P., 22:365-90
Dripps, R. D., 22:293-304
Düby, S., 22:305-26
Dupont, H. L., 24:103-10
Durie, B. G. M., 26:283-88

E

Edgington, T. S., 25:303-14
Ehrenkrantz, N. J., 25:9-14
Elliott, G. R., 26:499-520
Eriksson, M., 24:29-40

F

Fauci, A. S., 26:277-81
Fefer, A., 25:179-98
Feldman, D., 26:83-90
Fernstrom, J. D., 25:1-8
Field, J. B., 24:309-14
Finch, C. A., 22:221-34
Fisher, R. S., 26:373-90
Forget, B. G., 26:345-51
Formal, S. B., 24:103-10
Fox, A. C., 22:355-64
Fredrickson, D. S., 24:315-24
Freedman, D. X., 24:153-64
Friesen, H. G., 24:251-70
Fudenberg, H. H., 22:145-70; 24:175-208

G

Gabbay, K. H., 26:521-36
Gangarosa, E. J., 25:75-81
Gershon, S., 23:439-52
Gibbons, R. J., 26:121-36
Ginos, J., 22:305-26
Glassman, E., 23:229-44
Glenner, G. G., 25:131-35
Glucksberg, H., 25:179-98
Gold, P., 22:85-94
Goldberg, M., 26:537-50
Goldreyer, B. N., 26:219-28
Goldstein, G., 22:119-24
Good, R. A., 24:135-44
Gorman, R. E., 23:263-84
Gottlieb, L. S., 25:411-29
Gray, G. M., 22:391-404
Greenberg, P. L., 25:269-78
Guenter, C. A., 22:283-92
Gunnells, J. C. Jr., 26:259-75
Gunnar, R. M., 26:471-83

H

Hammarsten, J. F., 22:283-92
Harker, L. A., 25:383-400
Harmel, M. H., 26:137-41
Harrison, D. C., 25:143-48
Hellström, I., 23:19-38
Hellström, K. E., 23:19-38
Henderson, E. S., 26:289-306
Henkin, R. E., 26:193-201
Henkind, P., 22:95-102
Higby, D. J., 26:289-306
Hinman, F., 24:83-88
Hirschhorn, K., 24:67-74
Holick, M. F., 25:349-67
Holman, R. B., 26:499-520
Holmberg, L., 26:33-44
Hornick, R. B., 24:103-10
Horton, E. S., 22:235-50
Hough, P. V. C., 25:431-46
Hurvitz, A. I., 23:191-200
Hwang, P., 24:251-70

I

Ingbar, S. H., 26:443-49
Israel, C. W., 22:103-18

J

Jacquillat, C., 25:39-51
Janoff, A., 23:177-90
Johannes, R. S., 26:589-601
Johnson, J. W. C. Jr., 24:165-74
Johnson, M. T., 25:251-67
Jones, S. R., 25:251-67

K

Kalkhoff, R. K., 23:429-38
Kaltman, A. J., 22:343-54
Kane, J. J., 26:159-71
Karpatkin, S., 23:101-28
Kayden, H. J., 23:285-96
Kelly, J. F., 25:53-68
Kennedy, B. J., 26:235-43
Kessler, R. H., 23:365-84
Kessler, S., 24:89-102
Klebanoff, S. J., 22:39-62
Klerman, G. L., 25:199-217
Koenig, M. G., 22:25-38
Koffler, D., 25:149-64
Kolff, W. J., 23:321-32
Kwaan, H. C., 24:335-44

L

La Du, B. N., 23:453-68
Lane, J. M., 22:251-72
Lanning, C. F., 26:137-41
Lasseter, K. C., 25:69-74
Leader, R. W., 23:191-200
Legator, M. S., 23:413-28
Lenfant, C., 22:221-34
Le Quesne, L. P., 26:63-74
Lerner, K. G., 25:179-98
Levey, G. S., 25:69-74
Levin, A. S., 24:175-208
Levin, N. W., 24:365-84
Levy, R. I., 26:365-71
Lewy, P. R., 24:365-84
Lipinski, J., 26:551-65
Littlefield, J. W., 23:57-76
London, W. T., 23:161-76
Louria, D. B., 25:219-31
Lowenstein, J., 23:333-42
Luby, J. P., 25:251-67

M

Matthysse, S., 26:551-65
McCarty, D. J., 25:279-88
McConnell, R. B., 25:165-78
McCullough, J. L., 24:345-52
McGee, Z. A., 22:25-38
McGraw, C. P., 26:321-29
McGuffin, W. L. Jr., 26:259-75
McGuigan, J. E., 25:233-49
McGuire, W. L., 26:353-63
Menguy, R., 23:297-312, 313-20
Merimee, T. J., 25:137-42
Messner, R. P., 26:181-92
Millar, J. D., 22:251-72
Millman, I., 23:161-76
Mills, J. A., 22:185-220
Milunsky, A., 23:57-76
Mitchell, J. H., 25:369-81
Moos, R. H., 24:89-102
Moran, J. F., 26:471-83
Moss, G. S., 23:201-28
Murphy, D. L., 24:209-16
Murray, T., 26:537-50

N

Nathan, D. G., 26:345-51
Neer, R., 25:323-31
Neff, J. M., 22:251-72
Neill, C. A., 24:61-67
Neiman, P., 25:179-98
Nicoloff, D. M., 26:143-57
Nilsson, I. M., 26:33-44
Norman, P. S., 26:337-44

O

Odell, G. B., 24:165-74
Oken, D. E., 26:307-19
Oratz, M., 26:91-103

P

Palmer, R. F., 25:69-74
Papavasiliou, P. S., 22:305-26
Park, B. H., 24:135-44
Patterson, R., 25:53-68
Paulson, G. W., 26:75-81
Perkoff, G. T., 22:125-32
Perry, S., 22:171-84
Pink, R., 22:145-70
Pochi, P. E., 26:27-32
Pollack, R. E., 25:431-46
Potolsky, A., 24:75-82
Prien, E. L. Jr., 26:173-79

Q

Quinn, J. L. III, 26:193-201
Quintanilla, A., 24:365-84

R

Rabinowitz, M., 24:245-62
Rackley, C. E., 26:105-20
Raisfeld, I. H., 24:385-418
Ramot, B., 22:19-24
Ray, G. R., 26:567-88
Reiss, E., 24:217-32
Robertson, G. L., 25:315-22
Ross, G. T., 24:295-302
Rosse, W. F., 26:331-36
Rothschild, M. A., 26:91-103
Rubenstein, A. H., 22:1-8
Ruderman, N. B., 26:245-58
Russell, R. O. Jr., 26:105-20

S

Salans, L. B., 22:235-50
Salmon, S. E., 26:283-88
Sanford, J. P., 25:401-10, 26:485-98
Schaffner, W., 22:25-38
Schildkraut, J. J., 25:333-48
Schreiber, S. S., 26:91-103
Schrier, S. L., 25:269-78
Seegmiller, J. E., 25:15-28
Seligman, M. E. P., 23:407-12
Senay, E. C., 24:153-64
Sever, J. L., 26:589-601
Sharp, G. W. G., 24:19-28
Shoulson, I., 26:419-26
Sims, E. A. H., 22:235-50
Sinclair, J. C., 23:129-48
Smith, J. L., 22:103-18
Smith, R. B. III, 26:229-34
Smith, R. T., 23:39-56
Sokal, J. E., 24:281-88
Sokoloff, L., 24:271-80
Spencer, F. C., 23:229-44
Spitler, L. E., 24:175-208
Stamatoyannopoulos, G., 22:221-34
Stead, E. A. Jr., 22:273-82
Steck, A., 22:305-26
Steiner, D. F., 22:1-8
Storb, R., 25:179-98
Strauss, J. S., 26:27-32
Stutzman, L., 24:325-34
Sutnick, A. I., 23:161-76

T

Tan, C. T. C., 23:77-92
Tantum, K. R., 22:293-304
Terry, W. D., 25:131-35
Thomas, E. D., 25:179-98
Toole, J. F., 26:321-29
Trey, C., 25:411-29
Tschopp, T. B., 24:1-18
Twito, D. I., 26:235-43

U

Underwood, L. E., 26:427-41
Unger, R. D., 24:303-8
Utiger, R. D., 25:289-302

V

Vaitukaitis, J., 24:295-302
Valentine, W. N., 23:93-100
van Houte, J., 26:131-36
Van Wyk, J. J., 26:427-41
Ventura, A. K., 25:9-14
Vermylen, J., 25:447-55

Verstraete, M., 25:447-55

W

Wallace, W. A., 26:1-7
Wang, A.-C., 22:145-70
Warren, W. D., 26:229-34
Weil, M., 25:39-51
Weinstein, G. D., 24:345-52
Welch, M. H., 22:283-92
Wender, P. H., 23:355-72; 26:45-62
Wilber, J. F., 24:353-64
Wildenthal, K., 25:369-81
Wilhelm, D. L., 22:63-84
Williams, R. C. Jr., 26:181-92
Wilson, C. B., 25:83-98
Winick, M., 23:149-60
Winokur, G., 26:21-25
Wolfe, W. G., 26:203-17
Wolff, S. M., 26:277-81
Woodrow, J. C., 26:165-78

Y

Yaffe, S. J., 24:29-40
Yen, S. S. C., 26:403-17
Yu, P. N., 26:1-7

Z

Zak, R., 23:245-62
Zieve, L., 26:143-57
Zimmerman, T. S., 25:303-14
Zucker, M. B., 24:1-18

CHAPTER TITLES VOLUMES 22-26

ANESTHESIOLOGY		
Ketamine Anesthesia	C. F. Lanning, M. H. Harmel	26:137-41
Effects of Anesthesia on Intermediary Metabolism	E. A. Brunner, S. C. Cheng, M. L. Berman	26:391-401
CARDIOLOGY		
Papillary Muscle Dysfunction in Coronary (Ischemic) Heart Disease	N. P. DePasquale, G. E. Burch	22:327-42
Late Complications of Heart Valve Replacement	A. J. Kaltman	22:343-54
Abnormalities in Cardiac Contractility Associated with Localized Myocardial Damage	A. C. Fox	22:355-64
Structure and Function of the Myofibrillar Contractile Proteins	P. Dreizen	22:365-90
Pulmonary Involement in Hypovolemic Shock	G. S. Moss	23:201-28
Surgical Procedures for Coronary Artery Disease	F. C. Spencer, E. Glassman	23:229-44
Biochemical and Cellular Changes in Cardiac Hypertrophy	M. Rabinowitz, R. Zak	23:245-62
Genetics of Congenital Heart Disease	C. A. Neill	24:61-67
Viral Myocarditis and Its Sequelae	W. H. Abelmann	24:145-62
Primary Hyperlipoproteinemia in Infants	D. S. Fredrickson, J. L. Breslow	24:315-24
Sulfonylureas and the Heart	G. S. Levey, K. C. Lasseter, R. F. Palmer	25:69-74
Diphenylhydantoin as an Antiarrhythmic Drug	A. J. Atkinson, R. Davison	25:99-113
Antiarrhythmic Actions of Lidocaine	D. C. Harrison, K. A. Collinsworth	25:143-48
Static (Isometric) Exercise and the Heart: Physiological and Clinical Considerations	J. H. Mitchell, K. Wildenthal	25:369-81
Sudden Death and the Pre-Hospital Phase of Acute Myocardial Infarction	W. A. Wallace, P. N. Yu	26:1-7
Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis	L. P. Le Quesne	26:63-74
Left Ventricular Function in Acute and Chronic Artery Disease	C. E. Rackley, R. O. Russell Jr.	26:105-20
Clinical Pharmacology of Digitalis Glycosides	J. E. Doherty, J. J. Kane	26:159-71

Mechanisms of Supraventricular Tachycardias	B. N. Goldreyer	26:219-28
Selective Distal Splenorenal Shunt for Bleeding Esophageal Varices	R. B. Smith III, W. D. Warren	26:229-34
Low Renin Hypertension	J. C. Gunnells Jr., W. L. McGuffin Jr.	26:259-75
The Steal Syndromes	J. F. Toole, C. P. McGraw	26:321-29
The Total Correction of Congenital Heart Disease	E. Aberdeen	26:451-64
Pathophysiology of Heart Block	J. F. Moran, R. M. Gunnar	26:471-83
CLINICAL PHARMACOLOGY		
Drug Metabolism in the Newborn	M. Eriksson, S. J. Yaffe	24:29-40
Clinical Pharmacology of Drug Interactions	I. H. Raisfeld	24:385-418
Tardive Dyskinesia	G. W. Paulson	26:75-81
Clinical Pharmacology of Digitalis Glycosides	J. E. Doherty, J. J. Kane	26:159-71
Analgesic Abuse and Renal Disease	T. Murray, M. Goldberg	26:537-50
DERMATOLOGY		
Cytokinetics in Diseases of Epidermal Hyperplasia	G. D. Weinstein, J. L. McCullough	24:345-52
Skin Lipids and Acne	J. S. Strauss, P. E. Pochi, D. T. Downing	26:27-32
ENDOCRINOLOGY		
Proinsulin	A. H. Rubenstein, D. F. Steiner	22:1-18
Chemical Mediation of Hormone Action	M. W. Bitensky, R. E. Gorman	23:263-84
Abetalipoproteinemia	H. J. Kayden	23:285-96
Blood Levels of Parathyroid Hormone in Disorders of Calcium Metabolism	E. Reiss, J. M. Canterbury	24:217-32
Insulin Receptors	B. Desbuquois, P. Cuatrecasas	24:233-40
Human Prolactin	H. Friesen, P. Hwang	24:251-70
Recent Advances in Evaluation of Gonadotropic Hormones	J. L. Vaitukaitis, G. T. Ross	24:295-302
Normal and Abnormal Alpha Cell Function	R. H. Unger	24:303-8
Thyrotropin Releasing Hormone: Secretion and Actions	J. F. Wilber	24:353-64
Isolated Growth Hormone Deficiency and Related Disorders	T. J. Merimee	25:137-42
Serum Triiodothyronine in Man	R. D. Utiger	25:289-302
Vasopressin in Osmotic Regulation in Man	G. L. Robertson	25:315-22
Vitamin D Metabolism	M. F. Holick, H. F. DeLuca	25:349-67
Metabolic Response to Infection	W. R. Beisel	26:9-20
The Role of Hormone Receptors in the Action of Adrenal Steroids	D. Feldman	26:83-90
Regulation of Albumin Metabolism	M. A. Rothschild, M. Oratz, S. S. Schreiber	26:91-103
Muscle Amino Acid Metabolism and Gluconeogenesis	N. B. Ruderman	26:245-58
Low Renin Hypertension	J. C. Gunnells Jr., W. L. McGuffin Jr	26:259-75
Endocrine Therapy of Breast Cancer	W. L. McGuire	26:353-63
Interconversions of Apolipoprotein Fragments	C. B. Blum, R. I. Levy	26:365-71
Effects of Anesthesia on Intermediary Metabolism	E. A. Brunner, S. C. Cheng, M. L. Berman	26:391-401
Gonadotropin-Releasing Hormone	S. S. C. Yen	26:403-17
Relation Between Growth Hormone and Somatomedin	J. J. VanWyk, L. E. Underwood	26:427-41
Active Form of the Thyroid Hormone and Somatomedin	S. H. Ingbar, L. E. Braverman	26:443-49
Hyperglycemia, Polyol Metabolism, and the Complications of Diabetes Mellitus	K. H. Gabbay	26:521-36
GASTROENTEROLOGY		
Malabsorption Due to Lymphomatous Disease	B. Ramot	22:19-24
Intestinal Digestion and Maldigestion of Dietary Carbohydrates	G. M. Gray	22:391-404
Acute Gastric Mucosal Bleeding	R. Menguy	23:297-312
Choice of Operations in the Management of Symptomatic Sliding Hiatal Hernia	R. Menguy	23:313-20

Action of Cholera Toxin on Fluid and Electrolyte Movement in the Small Intestine	G. W. G. Sharp	24:19-28
Enterotoxic Diarrheal Syndromes	S. B. Formal, H. L. Dupont, R. B. Hornick	24:103-10
Extraction of Insulin by Liver	J. B. Field	24:309-14
Role of Gastrin in Hypersecretory Disorders in Man	J. E. McGuigan	25:233-40
Regulation of Albumin Metabolism	M. A. Rothschild, M. Oratz, S. S. Schreiber	26:91-103
Pathogenesis of Hepatic Coma	L. Zieve, D. M. Nicoloff	26:143-57
Selective Distal Splenorenal Shunt for Bleeding Esophageal Varices	R. B. Smith III, W. D. Warren	26:229-34
Disorders of the Lower Esophageal Sphincter	R. S. Fisher, S. Cohen	26:373-90

HEMATOLOGY

Proliferation of Myeloid Cells	S. Perry	22:171-84
Lymphocyte Physiology	J. A. Mills, S. R. Cooperband	22:185-220
Abnormal Hemoglobins with High and Low Oxygen Affinity	G. Stamatoyannopoulos, A. J. Bellingham, C. Lenfant, C. A. Finch	22:221-34
Treatment of Acute Lymphoblastic Leukemia	J. H. Burchenal, M. D. Dowling Jr., C. T. C. Tan	23:77-92
Red Cell Enzyme Deficiencies as a Cause of Hemolytic Disorders	W. N. Valentine	23:93-100
Human Platelet Senescence	S. Karpatkin	23:101-28
Platelet Function Tests in Disease	T. B. Tschopp, M. B. Zucker	24:1-18
Radiation and Drug Therapies, and Leukemia	A. Potolsky, W. P. Creger	24:75-82
Current Concepts in the Treatment of Chronic Myelocytic Leukemia	J. E. Sokal	24:281-88
Nature and Biological Activities of Degradation Products of Fibrinogen and Fibrin	H. C. Kwaan, G. H. Barlow	24:335-44
2, 3-DPG and Erythrocyte Oxygen Affinity	G. J. Brewer	25:29-38
Treatment of Acute Granulocytic Leukemias	J. Bernard, M. Weil, C. Jacquillat	25:39-51
Mechanism of Action of Hemoglobin	A. Arnone	25:123-30
Immuno-Prevention of Rh Hemolytic Disease of the Newborn	R. B. McConnell, J. C. Woodrow	25:165-78
Marrow Transplantation for Aplastic Anemia and Acute Leukemia	P. Neiman, E. D. Thomas, C. D. Buckner, R. Storb, A. Fefer, H. Glucksberg, K. G. Lerner	25:179-98
Clinical Utility of In Vitro Evaluation of Granulopoiesis	P. L. Greenberg, S. L. Schrier	25:269-78
Molecular Immunology of Factor VIII	T. S. Zimmerman, T. S. Edgington	25:303-14
Control of Platelet Production	L. A. Harker	25:383-400
Intravascular Coagulation in Liver Disease	M. Verstraete, J. Vermylen, D. Collen	25:447-55
Von Willebrand's Disease	L. Holmberg, I. M. Nilsson	26:33-44
Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis	L. P. Le Quesne	26:63-74
Cellular Kinetics, Staging, and Immunoglobulin Synthesis in Multiple Myeloma	B. G. M. Durie, S. E. Salmon	26:283-88
Granulocyte Transfusion Therapy	D. J. Higby, E. S. Henderson	26:289-306
The Antiglobulin Test in Autoimmune Hemolytic Anemia	W. F. Rosse	26:331-36
Thalassemia	B. G. Forget, D. G. Nathan	26:345-51

IMMUNOLOGY

Antibody Variability	R. Pink, A.-C. Wang, H. H. Fudenberg	22:145-70
Neutrophil Proteases in Inflammation	A. Janoff	23:177-90
Interspecies Patterns of Slow Virus Diseases	R. W. Leader, A. I. Hurvitz	23:191-200
Immunologic Reconsitution	W. D. Biggar, B. H. Park, R. A. Good	24:135-44
Transfer Factor Therapy in Immune Deficiency States	A. S. Levin, L. E. Spitler, H. H. Fudenberg	24:175-208

Animal Models of the Asthmatic State	R. Patterson, J. F. Kelly	25:53-68
Characterization of Amyloid	G. G. Glenner, W. D. Terry	25:131-35
Immunopathogenesis of Systemic Lupus Erythematosus	D. Koffler	25:149-64
Alterations in T- and B-Cells in Human Disease States	R. C. Williams Jr., R. P. Messner	26:181-92
The Antiglobulin Test in Autoimmune Hemolytic Anemia	W. F. Rosse	26:331-36
Immunotherapy (Desensitization) in Allergic Disease	P. S. Norman	26:337-44
INFECTIOUS DISEASE		
Superinfection in Lymphoreticular Diseases	Z. A. McGee, W. Schaffner, M. G. Koenig	22:25-38
The Role of Australia Antigen in Viral Hepatitis and Other Diseases	A. I. Sutnick, I. Millman, W. T. London, B. S. Blumberg	23:161-76
Preventive Antibiotic Management in Surgery	J. F. Burke	24:289-94
Venezuelan Equine Encephalitis Virus Infection in Man	N. J. Ehrenkrantz, A. K. Ventura	25:9-14
Food Poisoning Due to Vibrio Parahaemolyticus	W. H. Barker Jr., E. J. Gangarosa	25:75-81
Preventive Therapy of Tuberculosis	D. L. Brummer	25:115-22
Infectious Complications of Nonalcoholic Drug Abuse	D. B. Louria	25:219-31
Antiviral Chemotherapy	J. P. Luby, M. T. Johnson, S. R. Jones	25:251-67
Metabolic Response to Infection	W. R. Beisel	26:9-20
Granulocyte Transfusion Therapy	D. J. Higby, E. S. Henderson	26:289-306
Urinary Tract Symptoms and Infections	J. P. Sanford	26:485-98
MISCELLANEOUS		
Kinins in Human Disease	D. L. Wilhelm	22:63-84
Chromosome Identification	K. Hirschhorn	24:67-74
Genetic Aspects of Gout	M. A. Becker, J. E. Seegmiller	25:15-28
Medical Aspects of Recreational Skin and Scuba Diving	J. P. Sanford	25:401-10
Dental Caries	R. J. Gibbons, J. van Houte	26:121-36
Unusual Etiologies of Fever and Their Evaluation	S. M. Wolff, A. S. Fauci, D. C. Cole	26:277-81
MUSCULO-SKELETAL DISORDERS		
Myasthenia Gravis and the Thymus	G. Goldstein	22:119-24
Alcoholic Myopathy	G. T. Perkoff	22:125-32
Crystal Deposition Joint Disease	D. J. McCarty	25:279-88
NEOPLASIA		
Scanning Techniques to Assess Thyroid Nodules	J. L. Quinn III, R. E. Henkin	26:193-201
Treatment of Testicular Cancer	D. I. Twito, B. J. Kennedy	26:235-43
Cellular Kinetics, Staging and Immunoglobulin Synthesis in Multiple Myeloma	B. G. M. Durie, S. E. Salmon	26:283-88
Endocrine Therapy of Breast Cancer	W. L. McGuire	26:353-63
The Role of Radiation Therapy in the Definitive Treatment of Adenocarcinoma of the Prostate	G. R. Ray, M. A. Bagshaw	26:567-88
NEPHROLOGY		
Intraleukocytic Microbicidal Defects	S. J. Klebanoff	22:39-62
Hemodialysis in the Management of Renal Disease	W. J. Kolff	23:321-32
Renin Assay in Hypertensive Disease	J. Lowenstein	23:333-42
Urethrovesical Dysfunction and Infection	F. Hinman	24:83-88
Renal Energy Metabolism and Sodium Reabsorption	P. R. Lewy, A. Quintanilla, N. W. Levin, R. H. Kessler	24:365-84
Immunopathology and Glomerulonephritis	C. B. Wilson, F. J. Dixon	25:83-98
Calcium Oxalate Renal Stones	E. L. Prien Jr	26:173-79
Low Renin Hypertension	J. C. Gunnells Jr. W. L. McGuffin Jr	26:259-75
Acute Renal Failure (Vasomotor Nephropathy): Micropunture Studies of the Pathogenetic		

Mechanism	D. E. Oken	26:307-19
Urinary Tract Symptoms and Infections	J. P. Sanford	26:485-98
Analgesic Abuse and Renal Disease	T. Murray, M. Goldberg	26:537-50
NEUROLOGY		
Thermal Control in Premature Infants	J. C. Sinclair	23:129-48
Nutrition, Environment, and Behavioral Development	M. Winick, J. Coombs	23:149-60
The Lesch-Nyhan Sydrome	W. L. Nyhan	24:41-60
Metabolism of Ketone Bodies by the Brain	L. Sokoloff	24:271-80
The Minimal Brain Dysfunction Syndrome	P. H. Wender	26:45-62
Tardive Dyskinesia	G. W. Paulson	26:75-81
Pathogenesis of Hepatic Coma	L. Zieve, D. M. Nicoloff	26:143-57
The Steal Syndromes	J. F. Toole, C. P. McGraw	26:321-29
Huntington's Disease	I. Shoulson, T. N. Chase	26:419-26
A Definition of Cerebral Death	B. Boshes	26:465-70
Neuroregulators and Sleep Mechanisms	R. B. Holman, G. R. Elliott, J. D. Barchas	26:499-520
Subacute Sclerosing Panencephalitis	R. S. Johannes, J. L. Sever	26:589-601
NUTRITION		
Inducible Metabolic Abnormalities During Development of Obesity	E.A. H. Sims, E. S. Horton, L. B. Salans	22:235-50
Modification of Brain Serotonin by the Diet	J. D. Fernstrom	25:1-8
ONCOLOGY		
Antigenic Reversion in Human Cancer	P. Gold	22:85-94
Ablation Therapy for Hormone-Dependent Tumors	T. L. Dao	23:1-18
Immunity to Neuroblastomas and Melanomas	K. E. Hellström, I. Hellström	23:19-38
Epstein-Barr Virus Infection in Relation to Infectious Mononucleosis and Burkitt's Lymphoma	R. T. Smith, J. C. Bausher	23:39-56
Current Concepts in Treatment of Lymphomas with Special Reference to Hodgkin's Disease	L. Stutzman	24:325-34
Medical Management of the Hypercalcemia of Malignancy	L. J. Deftos, R. Neer	25:323-31
The Cell Surface and Malignant Transformation	R. E. Pollack, P. V. C. Hough	25:431-46
OPHTHALMOLOGY AND OTORHINOLARYNGOLOGY		
Senile Macular Degeneration	P. Henkind	22:95-102
Optic Atrophy and Neurosyphilis	J. L. Smith, C. W. Israel	22:103-18
PSYCHIATRY		
Biogenic Amines and Behavior	R. J. Baldessarini	23:343-54
Adopted Children and Their Families in the Evaluation of Nature-Nurture Interactions in the Schizophrenic Disorders	P. H. Wender	23:355-72
Genetic Analysis of Human Behavior	B. Childs	23:373-406
Learned Helplessness	M. E. P. Seligman	23:407-12
Chemical Mutagens	M. S. Legator	23:413-28
Behavioral Aspects of Chromosomal Disorders	S. Kessler, R. H. Moos	24:89-102
Mental Effects of L-Dopa	D. L. Murphy	24:209-16
Pharmacotherapy of Schizophrenia	G. L. Klerman	25:199-217
Biogenic Amines and Affective Disorders	J. J. Schildkraut	25:333-48
X-Linkage in Manic-Depressive Illness	R. J. Cadoret, G. Winokur	26:21-25
Biochemical Aspects of Schizophrenia	S. Matthysse, J. Lipinski	26:551-65
PUBLIC HEALTH AND ENVIRONMENTAL MEDICINE		
Smallpox and Smallpox Vaccination Policy	J. M. Lane, J. D. Millar, J. M. Neff	22:251-72
Use of Physicians' Assistants in the Delivery of Medical Care	E. A. Stead Jr.	22:273-82
Methadone Treatment of Heroin Addiction	D. X. Freedman, E. C. Senay	24:153-64
RADIOLOGY		
Diagnosis and Prevention of Postoperative Deep-Vein Thrombosis	L. P. Le Quesne	26:63-74
Scanning Techniques to Assess Thyroid Nodules	J. L. Quinn III, R. E. Henkin	26:193-201
The Role of Radiation Therapy in the Definitive Treatment of Adenocardinoma of the Prostate	G. R. Ray, M. A. Bagshaw	26:567-88
REPRODUCTIVE SYSTEM AND OBSTETRICS		
Fetal Indications for Therapeutic Abortion	A. C. Barnes	22:133-44

The Prenatal Diagnosis of Inborn Errors of Metabolism	A. Milunsky, J. W. Littlefield	23:57-76
Estimation of Gestational Age in the Fetal State	J. W. C. Johnson Jr., G. B. Odell	24:165-74
Gonadotropin-Releasing Hormone	S. S. C. Yen	26:403-17
RESPIRATORY SYSTEM		
Alpha1 Antitrypsin Deficiency and Pulmonary Emphysema	C. A. Guenter, M. H. Welch, J. F. Hammarsten	22:283-92
Clinical Application of Continuous Monitoring of Respiratory Failure	K. R. Tantum, R. D. Dripps	22:293-304
Mechanics of Air Flow in Airway Obstruction	A. C. DeGraff Jr., A. Bouhuys	24:111-34
Endogenous Carbon Monoxide Metabolism	R. F. Coburn	24:241-50
Oxygen Toxicity	W. G. Wolfe, W. C. DeVries	26:203-17
THERAPEUTICS AND TOXICOLOGY		
Metabolic Modification of Parkinson's Disease and of Chronic Manganese Poisoning	G. C. Cotzias, P. S. Papavasiliou, J. Ginos, A. Steck, S. Düby	22:305-26
Effects of Oral Contraceptive Agents and Sex Steroids on Carbohydrate Metabolism	R. K. Kalkhoff	23:429-38
Lithium Salts in the Management of the Manic-Depressive Syndrome	S. Gershon	23:439-52
Pharmacogenetics: Defective Enzymes in Relation to Reactions to Drugs	B. N. La Du	23:453-68
Drug Metabolism in the Newborn	M. Eriksson, S. J. Yaffe	24:29-40
Clinical Pharmacology of Drug Interactions	I. H. Raisfeld	24:385-418
The Effects of Fluorinated Anesthetics on the Liver and Kidneys	L. S. Gottlieb, C. Trey	25:411-29